Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

  http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:38:28 ; Search time 56.732 Seconds
(without alignments)
408.044 Million cell updates/sec

Title: US-10-680-349-42
Perfect score: 1462
Sequence: 1 MNYKKILVRSALISLMSILP.....ASVTLDVGYFGGEIGMRFTF 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1462	100.0	280	2	US-09-660-587-42	Sequence 42, Appl
2	1462	100.0	280	2	US-09-314-701-48	Sequence 48, Appl
3	1462	100.0	280	2	US-09-811-007A-42	Sequence 42, Appl
4	1462	100.0	280	2	US-10-314-639-48	Sequence 48, Appl
5	1462	100.0	280	2	US-10-059-964A-48	Sequence 48, Appl
6	1202.5	82.3	283	2	US-09-660-587-10	Sequence 10, Appl
7	1202.5	82.3	283	2	US-09-261-358A-10	Sequence 10, Appl
8	1202.5	82.3	283	2	US-09-201-458-6	Sequence 6, Appli
9	1202.5	82.3	283	2	US-09-314-701-4	Sequence 4, Appli
10	1202.5	82.3	283	2	US-09-811-007A-10	Sequence 10, Appl
11	1202.5	82.3	283	2	US-10-314-639-4	Sequence 4, Appli
12	1202.5	82.3	283	2	US-10-059-964A-4	Sequence 4, Appli
13	644.5	44.1	281	2	US-09-660-587-9	Sequence 9, Appli
14	644.5	44.1	281	2	US-09-261-358A-9	Sequence 9, Appli
15	644.5	44.1	281	2	US-09-201-458-5	Sequence 5, Appli
16	644.5	44.1	281	2	US-09-811-007A-9	Sequence 9, Appli
17	642.5	43.9	281	2	US-09-314-701-2	Sequence 2, Appli
18	642.5	43.9	281	2	US-10-314-639-2	Sequence 2, Appli
19	642.5	43.9	281	2	US-10-059-964A-2	Sequence 2, Appli
20	642.5	43.9	281	2	US-10-059-964A-67	Sequence 67, Appl
21	629.5	43.1	276	2	US-08-953-326-18	Sequence 18, Appl
22	629.5	43.1	276	2	US-09-553-662-18	Sequence 18, Appl
23	629.5	43.1	276	2	US-10-062-994-18	Sequence 18, Appl
24	621	42.5	288	2	US-09-314-701-32	Sequence 32, Appl
25	621	42.5	288	2	US-10-314-639-32	Sequence 32, Appl
26	621	42.5	288	2	US-10-059-964A-32	Sequence 32, Appl
27	620	42.4	286	2	US-08-953-326-15	Sequence 15, Appl
28	620	42.4	286	2	US-09-660-587-12	Sequence 12, Appl
29	620	42.4	286	2	US-09-261-358A-12	Sequence 12, Appl
30	620	42.4	286	2	US-09-201-458-8	Sequence 8, Appli
31	620	42.4	286	2	US-09-314-701-8	Sequence 8, Appli
32	620	42.4	286	2	US-09-553-662-15	Sequence 15, Appl
33	620	42.4	286	2	US-10-062-994-15	Sequence 15, Appl
34	620	42.4	286	2	US-09-811-007A-12	Sequence 12, Appl
35	620	42.4	286	2	US-10-314-639-8	Sequence 8, Appli
36	620	42.4	286	2	US-10-059-964A-8	Sequence 8, Appli
37	605	41.4	280	2	US-09-660-587-14	Sequence 14, Appl
38	605	41.4	280	2	US-09-261-358A-14	Sequence 14, Appl
39	605	41.4	280	2	US-09-201-458-10	Sequence 10, Appl
40	605	41.4	280	2	US-09-314-701-12	Sequence 12, Appl
41	605	41.4	280	2	US-09-811-007A-14	Sequence 14, Appl
42	605	41.4	280	2	US-10-314-639-12	Sequence 12, Appl
43	605	41.4	280	2	US-10-059-964A-12	Sequence 12, Appl
44	603	41.2	280	2	US-08-953-326-17	Sequence 17, Appl
45	603	41.2	280	2	US-09-553-662-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-660-587-42
; Sequence 42, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/660,587
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-660-587-42

Query Match 100.0%; Score 1462; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.1e-151;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 2

```
US-09-314-701-48
; Sequence 48, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT: Ohasi, No. 6544517io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION: Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/09/314,701
;  CURRENT FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-09-314-701-48

  Query Match             100.0%;  Score 1462;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-151;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 3
US-09-811-007A-42
; Sequence 42, Application US/09811007A
; Patent No. 6660269
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/09/811,007A
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-811-007A-42

  Query Match             100.0%;  Score 1462;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-151;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 4
US-10-314-639-48
; Sequence 48, Application US/10314639
; Patent No. 6893640
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT:  Ohasi, No. 6893640io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION:  Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/10/314,639
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US/09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-314-639-48

  Query Match             100.0%;  Score 1462;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-151;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 5
US-10-059-964A-48
; Sequence 48, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
;  APPLICANT: RIKIHISA, YASUKO
```

```
;  APPLICANT:  OHASHI, NORIO
;  TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
;  TITLE OF INVENTION:  CHAFFEENSIS
;  FILE REFERENCE: 22727-04109
;  CURRENT APPLICATION NUMBER: US/10/059,964A
;  CURRENT FILING DATE:  2002-01-28
;  PRIOR APPLICATION NUMBER: 09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: 60/100,843
;  PRIOR FILING DATE: 1998-09-18
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-059-964A-48

  Query Match             100.0%;  Score 1462;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-151;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 6
US-09-660-587-10
; Sequence 10, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
;  APPLICANT:  Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE:  D6152CIP2
;  CURRENT APPLICATION NUMBER: US/09/660,587
;  CURRENT FILING DATE:  2000-09-12
;  PRIOR APPLICATION NUMBER:  09/261,358
;  PRIOR FILING DATE:  1999-03-03
;  NUMBER OF SEQ ID NOS:  46
; SEQ ID NO 10
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-660-587-10

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              ||||||  | ||||||||||||||||||||  |     ||::|||||| ||||||||||||||

Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
              ||  ||||  |:||||||||||||| :  :| |   | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSGIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:|||||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||||  ||||||| ||:| |||||||:|:|||| 
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 7
US-09-261-358A-10
; Sequence 10, Application US/09261358A
; Patent No. 6403780
; GENERAL INFORMATION:
;  APPLICANT:  Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE:  D6152CIP
;  CURRENT APPLICATION NUMBER: US/09/261,358A
;  CURRENT FILING DATE:  1999-03-03
;  PRIOR APPLICATION NUMBER:  09/201,458
;  PRIOR FILING DATE:  1998-11-30
;  NUMBER OF SEQ ID NOS:  33
; SEQ ID NO 10
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-261-358A-10

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              ||||||  | ||||||||||||||||||||  |     ||::|||||| ||||||||||||||

Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
              ||  ||||  |:||||||||||||| :  :| |   | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSGIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:|||||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||||  ||||||| ||:| |||||||:|:|||| 
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 8
US-09-201-458-6
; Sequence 6, Application US/09201458A
; Patent No. 6458942
; GENERAL INFORMATION:
```

```
;  APPLICANT: Walker, David H.
;  APPLICANT: McBride, Jere W.
;  APPLICANT: Yu, Xue-Jie
;  TITLE OF INVENTION: 28-kDa Immunoreactive Protein Gene of Ehrlichia
;  TITLE OF INVENTION: canis and Uses Thereof
;  FILE REFERENCE: D6152
;  CURRENT APPLICATION NUMBER: US/09/201,458A
;  CURRENT FILING DATE: 1998-11-30
;  NUMBER OF SEQ ID NOS: 21
; SEQ ID NO 6
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-201-458-6

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
              || ||||  |:|||||||||||| :  :| |  | ::||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSGIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:|||| :||| ||||:||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 9
US-09-314-701-4
; Sequence 4, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT: Ohasi, No. 6544517io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION: Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/09/314,701
;  CURRENT FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-09-314-701-4

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
              || ||||  |:|||||||||||| :  :| |  | ::||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSGIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:|||| :||| ||||:||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 10
US-09-811-007A-10
; Sequence 10, Application US/09811007A
; Patent No. 6660269
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/09/811,007A
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-811-007A-10

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
              || ||||  |:|||||||||||| :  :| |  | ::||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSGIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:|||| :||| ||||:||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 11
US-10-314-639-4
; Sequence 4, Application US/10314639
; Patent No. 6893640
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT:  Ohasi, No. 68936401o
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION:  Chaffeensis
;  FILE REFERENCE: 22727/04021
```

```
;  CURRENT APPLICATION NUMBER: US/10/314,639
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US/09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-10-314-639-4

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:|||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 12
US-10-059-964A-4
; Sequence 4, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
;  APPLICANT: RIKIHISA, YASUKO
;  APPLICANT:  OHASHI, NORIO
;  TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
;  TITLE OF INVENTION:  CHAFFEENSIS
;  FILE REFERENCE: 22727-04109
;  CURRENT APPLICATION NUMBER: US/10/059,964A
;  CURRENT FILING DATE:  2002-01-28
;  PRIOR APPLICATION NUMBER: 09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: 60/100,843
;  PRIOR FILING DATE: 1998-09-18
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-10-059-964A-4

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 5.4e-123;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:|||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 13
US-09-660-587-9
; Sequence 9, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
;  APPLICANT:  Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE:  D6152CIP2
;  CURRENT APPLICATION NUMBER: US/09/660,587
;  CURRENT FILING DATE:  2000-09-12
;  PRIOR APPLICATION NUMBER:  09/261,358
;  PRIOR FILING DATE:  1999-03-03
;  NUMBER OF SEQ ID NOS:  46
; SEQ ID NO 9
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis P28
US-09-660-587-9

  Query Match             44.1%;  Score 644.5;  DB 2;  Length 281;
  Best Local Similarity   48.1%;  Pred. No. 6.4e-62;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:: |       :  | ||       ::||   ||:|:||||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||     | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
               |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              |::| ||||:|  || | |     | :   | | ||| :|| |:| || |
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFAF 281


RESULT 14
US-09-261-358A-9
; Sequence 9, Application US/09261358A
; Patent No. 6403780
; GENERAL INFORMATION:
;  APPLICANT:  Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE:  D6152CIP
```

```
;  CURRENT APPLICATION NUMBER: US/09/261,358A
;  CURRENT FILING DATE:  1999-03-03
;  PRIOR APPLICATION NUMBER:  09/201,458
;  PRIOR FILING DATE:  1998-11-30
;  NUMBER OF SEQ ID NOS:  33
; SEQ ID NO 9
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis P28
US-09-261-358A-9

  Query Match             44.1%;  Score 644.5;  DB 2;  Length 281;
  Best Local Similarity   48.1%;  Pred. No. 6.4e-62;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:: |       :  | ||       ::||   ||:|:|||||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||     | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
               |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              |::| ||||:|  || | |     | :   | | ||| :|| |:| || |
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFAF 281


RESULT 15
US-09-201-458-5
; Sequence 5, Application US/09201458A
; Patent No. 6458942
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT: McBride, Jere W.
;  APPLICANT: Yu, Xue-Jie
;  TITLE OF INVENTION: 28-kDa Immunoreactive Protein Gene of Ehrlichia
;  TITLE OF INVENTION: canis and Uses Thereof
;  FILE REFERENCE: D6152
;  CURRENT APPLICATION NUMBER: US/09/201,458A
;  CURRENT FILING DATE: 1998-11-30
;  NUMBER OF SEQ ID NOS: 21
; SEQ ID NO 5
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis P28
US-09-201-458-5

  Query Match             44.1%;  Score 644.5;  DB 2;  Length 281;
  Best Local Similarity   48.1%;  Pred. No. 6.4e-62;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:: |       :  | ||       ::||   ||:|:|||||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||     | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
               |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              |::| ||||:|  || | |     | :   | | ||| :|| |:| || |
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFAF 281


Search completed: January 24, 2006, 11:49:00
Job time : 57.732 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:46:59 ; Search time 154.641 Seconds
(without alignments)
756.543 Million cell updates/sec

Title: US-10-680-349-42
Perfect score: 1462
Sequence: 1 MNYKKILVRSALISLMSILP.....ASVTLDVGYFGGEIGMRFTF 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1462	100.0	280	3	US-09-811-007-42	Sequence 42, Appl
2	1462	100.0	280	4	US-10-062-624-42	Sequence 42, Appl
3	1462	100.0	280	4	US-10-059-964-48	Sequence 48, Appl
4	1462	100.0	280	4	US-10-062-051-42	Sequence 42, Appl
5	1462	100.0	280	4	US-10-062-920-42	Sequence 42, Appl
6	1462	100.0	280	4	US-10-314-639-48	Sequence 48, Appl
7	1462	100.0	280	4	US-10-680-349-42	Sequence 42, Appl
8	1462	100.0	280	5	US-10-731-554-42	Sequence 42, Appl
9	1462	100.0	280	5	US-10-901-714-48	Sequence 48, Appl
10	1462	100.0	280	5	US-10-901-774-48	Sequence 48, Appl
11	1202.5	82.3	283	3	US-09-846-808-14	Sequence 14, Appl
12	1202.5	82.3	283	3	US-09-811-007-10	Sequence 10, Appl
13	1202.5	82.3	283	4	US-10-062-624-10	Sequence 10, Appl
14	1202.5	82.3	283	4	US-10-059-964-4	Sequence 4, Appli
15	1202.5	82.3	283	4	US-10-062-051-10	Sequence 10, Appl
16	1202.5	82.3	283	4	US-10-284-986-14	Sequence 14, Appl
17	1202.5	82.3	283	4	US-10-062-920-10	Sequence 10, Appl
18	1202.5	82.3	283	4	US-10-314-639-4	Sequence 4, Appli
19	1202.5	82.3	283	4	US-10-369-293-14	Sequence 14, Appl
20	1202.5	82.3	283	4	US-10-285-042-14	Sequence 14, Appl
21	1202.5	82.3	283	4	US-10-680-349-10	Sequence 10, Appl
22	1202.5	82.3	283	5	US-10-731-554-10	Sequence 10, Appl
23	1202.5	82.3	283	5	US-10-901-714-4	Sequence 4, Appli
24	1202.5	82.3	283	5	US-10-901-774-4	Sequence 4, Appli
25	644.5	44.1	281	3	US-09-846-808-19	Sequence 19, Appl
26	644.5	44.1	281	3	US-09-811-007-9	Sequence 9, Appli
27	644.5	44.1	281	4	US-10-062-624-9	Sequence 9, Appli
28	644.5	44.1	281	4	US-10-062-051-9	Sequence 9, Appli
29	644.5	44.1	281	4	US-10-284-986-19	Sequence 19, Appl
30	644.5	44.1	281	4	US-10-062-920-9	Sequence 9, Appli
31	644.5	44.1	281	4	US-10-369-293-19	Sequence 19, Appl
32	644.5	44.1	281	4	US-10-285-042-19	Sequence 19, Appl
33	644.5	44.1	281	4	US-10-680-349-9	Sequence 9, Appli
34	644.5	44.1	281	5	US-10-731-554-9	Sequence 9, Appli
35	642.5	43.9	281	4	US-10-059-964-2	Sequence 2, Appli
36	642.5	43.9	281	4	US-10-314-639-2	Sequence 2, Appli
37	642.5	43.9	281	5	US-10-901-714-2	Sequence 2, Appli
38	642.5	43.9	281	5	US-10-901-714-67	Sequence 67, Appl
39	642.5	43.9	281	5	US-10-901-774-2	Sequence 2, Appli
40	642.5	43.9	281	5	US-10-901-774-67	Sequence 67, Appl
41	629.5	43.1	276	4	US-10-062-994-18	Sequence 18, Appl
42	629.5	43.1	276	4	US-10-062-994-18	Sequence 18, Appl
43	629.5	43.1	276	4	US-10-722-077-18	Sequence 18, Appl
44	621	42.5	288	4	US-10-059-964-32	Sequence 32, Appl
45	621	42.5	288	4	US-10-314-639-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-811-007-42
; Sequence 42, Application US/09811007
; Publication No. US20030185849A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-811-007-42

Query Match 100.0%; Score 1462; DB 3; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-136;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 2

US-10-062-624-42
; Sequence 42, Application US/10062624
; Publication No. US20020115840A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/062,624
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-062-624-42

Query Match 100.0%; Score 1462; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-136;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 3
US-10-059-964-48
; Sequence 48, Application US/10059964
; Publication No. US20020120115A1
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. US20020120115A1io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/10/059,964
; CURRENT FILING DATE: 2002-01-28
; EARLIER APPLICATION NUMBER: 09/314,701
; EARLIER FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
US-10-059-964-48

Query Match 100.0%; Score 1462; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-136;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 4
US-10-062-051-42
; Sequence 42, Application US/10062051
; Publication No. US20030073095A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/062,051
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/660,587
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-062-051-42

Query Match 100.0%; Score 1462; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-136;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

```
RESULT 5
US-10-062-920-42
; Sequence 42, Application US/10062920
; Publication No. US20030096250A1
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/10/062,920
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US/09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 09/261,358
;  PRIOR FILING DATE: 1999-03-03
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-062-920-42

  Query Match           100.0%;  Score 1462;  DB 4;  Length 280;
  Best Local Similarity  100.0%;  Pred. No. 9.8e-136;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 6
US-10-314-639-48
; Sequence 48, Application US/10314639
; Publication No. US20030103991A1
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT:  Ohasi, No. US20030103991A1io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION:  Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/10/314,639
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US/09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-314-639-48

  Query Match           100.0%;  Score 1462;  DB 4;  Length 280;
  Best Local Similarity  100.0%;  Pred. No. 9.8e-136;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 7
US-10-680-349-42
; Sequence 42, Application US/10680349
; Publication No. US20040198951A1
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2/D1
;  CURRENT APPLICATION NUMBER: US/10/680,349
;  CURRENT FILING DATE:  2003-10-07
;  PRIOR APPLICATION NUMBER: US/10/062,624
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: 09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-680-349-42

  Query Match           100.0%;  Score 1462;  DB 4;  Length 280;
  Best Local Similarity  100.0%;  Pred. No. 9.8e-136;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 8
US-10-731-554-42
; Sequence 42, Application US/10731554
; Publication No. US20040247616A1
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/10/731,554
;  CURRENT FILING DATE:  2003-12-09
;  PRIOR APPLICATION NUMBER: US/09/811,007
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-731-554-42

  Query Match             100.0%;  Score 1462;  DB 5;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-136;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 9
US-10-901-714-48
; Sequence 48, Application US/10901714
; Publication No. US20040265333A1
; GENERAL INFORMATION:
;  APPLICANT: RIKIHISA, YASUKO
;  APPLICANT:  OHASHI, NORIO
;  TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
;  TITLE OF INVENTION:  CHAFFEENSIS
;  FILE REFERENCE: 22727-04109
;  CURRENT APPLICATION NUMBER: US/10/901,714
;  CURRENT FILING DATE:  2004-07-29
;  PRIOR APPLICATION NUMBER: 09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: 60/100,843
;  PRIOR FILING DATE: 1998-09-18
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-901-714-48

  Query Match             100.0%;  Score 1462;  DB 5;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-136;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 10
US-10-901-774-48
; Sequence 48, Application US/10901774
; Publication No. US20040265334A1
; GENERAL INFORMATION:
;  APPLICANT: RIKIHISA, YASUKO
;  APPLICANT:  OHASHI, NORIO
;  TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
;  TITLE OF INVENTION:  CHAFFEENSIS
;  FILE REFERENCE: 22727-04109
;  CURRENT APPLICATION NUMBER: US/10/901,774
;  CURRENT FILING DATE:  2004-07-29
;  PRIOR APPLICATION NUMBER: 09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: 60/100,843
;  PRIOR FILING DATE: 1998-09-18
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-901-774-48

  Query Match             100.0%;  Score 1462;  DB 5;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-136;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
```

```
Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 11
US-09-846-808-14
; Sequence 14, Application US/09846808
; Patent No. US20020064531A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: Yu, Xu-Jie
; TITLE OF INVENTION: Ehrlichia chaffeensis 28 kDa Outer Membrane
; TITLE OF INVENTION: Protein Multigene Family
; FILE REFERENCE: D6311
; CURRENT APPLICATION NUMBER: US/09/846,808
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,035
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 14
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: P28-14 Outer Membrane Protein of
; OTHER INFORMATION: Ehrlichia chaffeensis
US-09-846-808-14

Query Match 82.3%; Score 1202.5; DB 3; Length 283;
Best Local Similarity 79.5%; Pred. No. 4.4e-110;
Matches 225; Conservative 26; Mismatches 29; Indels 3; Gaps 1;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| |||:|||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283
```

RESULT 12
US-09-811-007-10
; Sequence 10, Application US/09811007
; Publication No. US20030185849A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-811-007-10

Query Match 82.3%; Score 1202.5; DB 3; Length 283;
Best Local Similarity 79.5%; Pred. No. 4.4e-110;
Matches 225; Conservative 26; Mismatches 29; Indels 3; Gaps 1;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| |||:|||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283
```

RESULT 13
US-10-062-624-10
; Sequence 10, Application US/10062624
; Publication No. US20020115840A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/062,624
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-10-062-624-10

Query Match 82.3%; Score 1202.5; DB 4; Length 283;
Best Local Similarity 79.5%; Pred. No. 4.4e-110;
Matches 225; Conservative 26; Mismatches 29; Indels 3; Gaps 1;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| |||:|||||||||||||||||||||||||||||
```

```
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 14
US-10-059-964-4
; Sequence 4, Application US/10059964
; Publication No. US20020120115A1
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT: Ohasi, No. US20020120115A1io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION: Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/10/059,964
;  CURRENT FILING DATE: 2002-01-28
;  EARLIER APPLICATION NUMBER: 09/314,701
;  EARLIER FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-10-059-964-4

  Query Match             82.3%;  Score 1202.5;  DB 4;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 4.4e-110;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:|||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| |||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 15
US-10-062-051-10
; Sequence 10, Application US/10062051
; Publication No. US20030073095A1
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/10/062,051
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US/09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 09/261,358
;  PRIOR FILING DATE: 1999-03-03
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-10-062-051-10

  Query Match             82.3%;  Score 1202.5;  DB 4;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 4.4e-110;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:|||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| |||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


Search completed: January 24, 2006, 12:03:33
Job time : 155.641 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:47:54 ; Search time 31.1111 Seconds
(without alignments)
91.208 Million cell updates/sec

Title: US-10-680-349-42
Perfect score: 1462
Sequence: 1 MNYKKILVRSALISLMSILP.....ASVTLDVGYFGGEIGMRFTF 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	93.5	6.4	658	6	US-10-873-528-17	Sequence 17, Appl
2	93.5	6.4	677	6	US-10-873-528-155	Sequence 155, App
3	89.5	6.1	1288	7	US-11-052-554A-93	Sequence 93, Appl
4	87.5	6.0	669	6	US-10-997-201A-30	Sequence 30, Appl
5	87.5	6.0	1255	7	US-11-052-554A-271	Sequence 271, App
6	87.5	6.0	1255	7	US-11-052-554A-272	Sequence 272, App
7	86	5.9	657	6	US-10-957-880-4	Sequence 4, Appli
8	86	5.9	1255	7	US-11-022-562-235	Sequence 235, App
9	86	5.9	1255	7	US-11-052-554A-265	Sequence 265, App
10	86	5.9	1255	7	US-11-052-554A-266	Sequence 266, App
11	86	5.9	1255	7	US-11-052-554A-267	Sequence 267, App
12	86	5.9	1255	7	US-11-052-554A-268	Sequence 268, App
13	86	5.9	1255	7	US-11-052-554A-269	Sequence 269, App
14	86	5.9	1255	7	US-11-052-554A-270	Sequence 270, App
15	86	5.9	1255	7	US-11-052-554A-273	Sequence 273, App
16	86	5.9	1255	7	US-11-052-554A-274	Sequence 274, App
17	86	5.9	1255	7	US-11-052-554A-275	Sequence 275, App
18	86	5.9	1279	6	US-10-957-880-3	Sequence 3, Appli
19	84	5.7	1123	7	US-11-037-243-77	Sequence 77, Appl
20	81.5	5.6	628	7	US-11-080-991-108	Sequence 108, App
21	79.5	5.4	1296	6	US-10-615-668-3	Sequence 3, Appli
22	79.5	5.4	2367	7	US-11-051-453-42	Sequence 42, Appl
23	79	5.4	1458	7	US-11-096-274-2	Sequence 2, Appli
24	78.5	5.4	423	7	US-11-167-856-8	Sequence 8, Appli
25	78.5	5.4	2516	6	US-10-647-956A-2	Sequence 2, Appli
26	78	5.3	229	6	US-10-467-657-6238	Sequence 6238, Ap
27	78	5.3	1034	6	US-10-392-234A-30	Sequence 30, Appl
28	77.5	5.3	585	7	US-11-108-172-1067	Sequence 1067, Ap
29	77	5.3	390	7	US-11-052-554A-221	Sequence 221, App
30	77	5.3	1579	7	US-11-052-554A-9	Sequence 9, Appli
31	76.5	5.2	3063	7	US-11-186-284-26	Sequence 26, Appl
32	76.5	5.2	3194	7	US-11-052-554A-90	Sequence 90, Appl
33	76	5.2	255	6	US-10-510-386-66	Sequence 66, Appl
34	76	5.2	255	6	US-10-510-386-156	Sequence 156, App
35	76	5.2	2710	7	US-11-051-453-41	Sequence 41, Appl
36	75.5	5.2	444	7	US-11-074-176-170	Sequence 170, App
37	75.5	5.2	481	6	US-10-467-657-4660	Sequence 4660, Ap
38	75.5	5.2	1389	6	US-10-467-657-334	Sequence 334, App
39	75.5	5.2	2053	7	US-11-013-759-9	Sequence 9, Appli
40	75	5.1	329	7	US-11-052-554A-246	Sequence 246, App
41	75	5.1	809	6	US-10-467-657-1106	Sequence 1106, Ap
42	75	5.1	1038	7	US-11-107-028-1	Sequence 1, Appli
43	74.5	5.1	482	6	US-10-878-556A-139	Sequence 139, App
44	74.5	5.1	1119	6	US-10-131-826A-352	Sequence 352, App
45	74.5	5.1	1122	7	US-11-191-374-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-873-528-17
; Sequence 17, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe PG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/10/873,528
; CURRENT FILING DATE: 2004-06-23
; PRIOR APPLICATION NUMBER: US/09/769,787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-873-528-17

Query Match 6.4%; Score 93.5; DB 6; Length 658;
Best Local Similarity 22.5%; Pred. No. 0.86;
Matches 47; Conservative 33; Mismatches 92; Indels 37; Gaps 10;

```
Qy          73 FGLKKDGDITKK----DDFTRV-----APGIDFQNNLISGFS-GSIGYSMDGPRIELEAA 122
               | || || | :|    |  ::      : |   :|  : |:  || |  : |       ||
Db         328 FYLKSDGKIAEKEWVYDSHSQAWYYFKSGGYMAKNETVDGYQLGSDGKWLGGKTTNENAA 387

Qy         123 YQQFNPKNTDNNDTDNGEYYKHFALSR-----KDAMEDQQYVVLKNDGIT-FMSLMVNTC 176
               | |  |   :  |:| ||   : :          ||   | : : :   |:: :|
Db         388 YYQVVPVTANVYDSD-GEKLSYISQGSVVWLDKDRKSDDKRLAITISGLSGYMK-----T 441

Qy         177 YDITAEGVS--FVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAF-IGGYY 233
               |: |   |  |:||           ::    :| :       | |:   :|  :|  |
Db         442 EDLQALDASKDFIPY-----------YESDGHRFYHYVAQNASIPVASHLSDMEVGKKY 489

Qy         234 HGVIGNKFEKIPVITPVVLNDAPQTTSAS 262
               :   |  |:   :  | :  |  : |: |
Db         490 YSADGLHFDGFKLENPFLFKDLTEATNYS 518
```

RESULT 2

```
US-10-873-528-155
; Sequence 155, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
;  APPLICANT: Microbial Technics Limited
;  APPLICANT:  Gilbert, Christophe FG
;  APPLICANT:  Hansbro, Philip M
;  TITLE OF INVENTION: Proteins
;  FILE REFERENCE: PWC/P21129WO
;  CURRENT APPLICATION NUMBER: US/10/873,528
;  CURRENT FILING DATE:  2004-06-23
;  PRIOR APPLICATION NUMBER: US/09/769,787
;  PRIOR FILING DATE: 2001-01-26
;  PRIOR APPLICATION NUMBER: GB 9816337.1
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: US 60/125164
;  PRIOR FILING DATE: 1999-03-19
;  NUMBER OF SEQ ID NOS: 388
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 155
;   LENGTH: 677
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-10-873-528-155

  Query Match             6.4%;  Score 93.5;  DB 6;  Length 677;
  Best Local Similarity   22.5%;  Pred. No. 0.89;
  Matches   47;  Conservative   33;  Mismatches   92;  Indels   37;  Gaps   10;

Qy          73 FGLKKDGDITKK----DDFTRV-----APGIDFQNNLISGFS-GSIGYSMDGPRIELEAA 122
               | || || | :|    |  ::      : |   :|  : |:  || |  : |       ||
Db         347 FYLKSDGKIAEKEWVYDSHSQAWYYFKSGGYMAKNETVDGYQLGSDGKWLGGKTTNENAA 406

Qy         123 YQQFNPKNTDNNDTDNGEYYKHFALSR-----KDAMEDQQYVVLKNDGIT-FMSLMVNTC 176
               | |   |   :  |:| ||   : :         ||   | : : :   |:: :|
Db         407 YYQVVPVPVTANVYDSD-GEKLSYISQGSVVWLDKDRKSDDKRLAITISGLSGYMK-----T 460

Qy         177 YDITAEGVS--FVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAF-IGGYY 233
               |: |   |  |:||            ::    :| :      | |:   :|   :|  |
Db         461 EDLQALDASKDFIPY-----------YESDGHRFYHYVAQNASIPVASHLSDMEVGKKY 508

Qy         234 HGVIGNKFEKIPVITPVVLNDAPQTTSAS 262
               :   |  |:   :  | :  |  : |: |
Db         509 YSADGLHFDGFKLENPFLFKDLTEATNYS 537


RESULT 3
US-11-052-554A-93
; Sequence 93, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 93
;   LENGTH: 1288
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori J99
US-11-052-554A-93

  Query Match             6.1%;  Score 89.5;  DB 7;  Length 1288;
  Best Local Similarity   22.4%;  Pred. No. 4.9;
  Matches   54;  Conservative   39;  Mismatches   75;  Indels   73;  Gaps   13;

Qy           1 MNYKKILVRSALISL-------------MSILPYQSFADPVGSRTNDN---------- 35
               :|  |::  |  :::               :|:  |  |:: :||::  |
Db         529 VNINKLITASTNVAVKNFNINELIVKTNGISVGEYTHFSEDIGSQSRINTVRLETGTRSI 588

Qy          36 -------KEG-------FYIS------AKYNPSISHFRKFSAEETPIN--GTNSLTKKVF 73
                     | |       || |       |:   ::   ||| |  || |  ||: |
Db         589 FSGGVKFKSGEKLVINDFYYSPWNYFDARNVKNVEITRKF-ASSTPENPWGTSKLMFNNL 647

Qy          74 GLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIEL----EAAYQQFNPK 129
               | ::  :     |: :    || ||       |:| | : | ::       ||   |
Db         648 TLGQNA-VMDYSQFSNLTIQGDFINN-----QGTINYLVRGGKVATLNVGNAAAMMF--- 698

Qy         130 NTDNNDTDNGE-YYKHFAL--SRKDAMEDQQYVVLKNDGITFMSLMVNTCYDITAEGVSF 186
                  ||| |:   :||       | :| ::: ::|:||   | : ::   |       |:|
Db         699 ---NNDIDSATGFYKPLIKINSAQDLIKNTEHVLLKAKIIGYGNVSTGT------NGISN 749

Qy         187 V 187
               |
Db         750 V 750


RESULT 4
US-10-997-201A-30
; Sequence 30, Application US/10997201A
; Publication No. US20050249739A1
; GENERAL INFORMATION:
;  APPLICANT: Marasco, Wayne
;  APPLICANT:  Sui, Jianhua
;  TITLE OF INVENTION: Antibodies Against SARS-COV and Methods of Use Thereof
;  FILE REFERENCE: 20363-026
;  CURRENT APPLICATION NUMBER: US/10/997,201A
;  CURRENT FILING DATE:  2004-11-24
;  PRIOR APPLICATION NUMBER: 60/524,840
;  PRIOR FILING DATE: 2003-11-25
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
;   LENGTH: 669
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-997-201A-30

  Query Match             6.0%;  Score 87.5;  DB 6;  Length 669;
  Best Local Similarity   20.6%;  Pred. No. 3;
  Matches   54;  Conservative   33;  Mismatches  112;  Indels   63;  Gaps   13;

Qy          42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
               |  || :  | |: :    | :|  :    |:    |  : | ||  ::|||      |
Db         349 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 403

Qy          96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
               :   |   | | :                 :| :| |   | |  | |::    :   |
Db         404 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 448

Qy         155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
                |   |    ||        :| ||          |  |: : ||    :  :|:
Db         449 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 507

Qy         201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF---------IGGYYHGVIGNKFEKI 244
                    : |:  : | : |  |    :||    |         :   :    |  | :|
Db         508 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEI 566

Qy         245 PVITPVVLND----APQTTSAS 262
                 |:|          | | ::|
Db         567 LDISPCAFGGVSVITPGTNASS 588


RESULT 5
US-11-052-554A-271
; Sequence 271, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
```

```
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 271
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus
US-11-052-554A-271

  Query Match             6.0%;  Score 87.5;  DB 7;  Length 1255;
  Best Local Similarity   20.6%;  Pred. No. 7.1;
  Matches   54;  Conservative   33;  Mismatches  112;  Indels   63;  Gaps   13;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||     |
Db        353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                  :| :| |   | |  | |::   :    |
Db        408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF---------IGGYYHGVIGNKFEKI 244
                    : |:  : | : |  |    :||    |         :  :   |   | :|
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEI 570

Qy        245 PVITPVVLND----APQTTSAS 262
                |:|          | | ::|
Db        571 LDISPCAFGGVSVITPGTNASS 592


RESULT 6
US-11-052-554A-272
; Sequence 272, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 272
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus Tor2
US-11-052-554A-272

  Query Match             6.0%;  Score 87.5;  DB 7;  Length 1255;
  Best Local Similarity   20.6%;  Pred. No. 7.1;
  Matches   54;  Conservative   33;  Mismatches  112;  Indels   63;  Gaps   13;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||     |
Db        353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                  :| :| |   | |  | |::   :    |
Db        408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF---------IGGYYHGVIGNKFEKI 244
                    : |:  : | : |  |    :||    |         :  :   |   | :|
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEI 570

Qy        245 PVITPVVLND----APQTTSAS 262
                |:|          | | ::|
Db        571 LDISPCAFGGVSVITPGTNASS 592


RESULT 7
US-10-957-880-4
; Sequence 4, Application US/10957880
; Publication No. US20050282154A1
; GENERAL INFORMATION:
;  APPLICANT: Brigham & Women's Hospital, Inc.
;  APPLICANT:  Farzan, Michael R
;  APPLICANT:  Li, Wenhui
;  APPLICANT:  Moore, Michael J
;  TITLE OF INVENTION: Angiotensin-Converting Enzyme-2 as a Receptor for the SARS Coronav
;  FILE REFERENCE: 7570/80644
;  CURRENT APPLICATION NUMBER: US/10/957,880
;  CURRENT FILING DATE:  2004-10-05
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: SARS coronovirus
US-10-957-880-4

  Query Match             5.9%;  Score 86;  DB 6;  Length 657;
  Best Local Similarity   21.4%;  Pred. No. 4;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||     |
Db        356 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 410

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                  :| :| |   | |  | |::   :    |
Db        411 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 455

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        456 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 514

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                    : |:  : | : |  |    :||    |
Db        515 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 548


RESULT 8
US-11-022-562-235
; Sequence 235, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
;  APPLICANT: Ruprecht, Ruth M.
;  APPLICANT:  Shisong, Jiang
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
;  TITLE OF INVENTION:  A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
;  FILE REFERENCE: DFN-043CN
;  CURRENT APPLICATION NUMBER: US/11/022,562
;  CURRENT FILING DATE:  2004-12-22
;  PRIOR APPLICATION NUMBER: PCT/US03/20322
;  PRIOR FILING DATE: 2003-06-27
```

```
;  PRIOR APPLICATION NUMBER: 60/392718
;  PRIOR FILING DATE: 2002-06-27
;  NUMBER OF SEQ ID NOS: 340
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 235
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: Human Immunodeficiency Virus
US-11-022-562-235

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db        353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                  :| :| |   | | | |::    :    |
Db        408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                    : |:  : | : |  |    :||    |
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 9
US-11-052-554A-265
; Sequence 265, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 265
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus Frankfurt 1
US-11-052-554A-265

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db        353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                  :| :| |   | | | |::    :    |
Db        408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                    : |:  : | : |  |    :||    |
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 10
US-11-052-554A-266
; Sequence 266, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 266
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus HSR1
US-11-052-554A-266

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db        353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                  :| :| |   | | | |::    :    |
Db        408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                    : |:  : | : |  |    :||    |
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 11
US-11-052-554A-267
; Sequence 267, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 267
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus ZJ01
US-11-052-554A-267

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;
```

```
Qy          42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
               |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db         353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy          96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
               :   |   | | :                  :| :| |   | |  | |::    :    |
Db         408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy         155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
                |   |    ||       :| ||         |  |: : ||    :  :|:
Db         453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy         201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                     : |:  : | : |  |    :||    |
Db         512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 12
US-11-052-554A-268
; Sequence 268, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 268
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus TW1
US-11-052-554A-268

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy          42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
               |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db         353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy          96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
               :   |   | | :                  :| :| |   | |  | |::    :    |
Db         408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy         155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
                |   |    ||       :| ||         |  |: : ||    :  :|:
Db         453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy         201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                     : |:  : | : |  |    :||    |
Db         512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 13
US-11-052-554A-269
; Sequence 269, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 269
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus CUHK-Su10
US-11-052-554A-269

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy          42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
               |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db         353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy          96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
               :   |   | | :                  :| :| |   | |  | |::    :    |
Db         408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy         155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
                |   |    ||       :| ||         |  |: : ||    :  :|:
Db         453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy         201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                     : |:  : | : |  |    :||    |
Db         512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 14
US-11-052-554A-270
; Sequence 270, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 270
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus Urbani
US-11-052-554A-270

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy          42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
               |  ||  :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db         353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy          96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
               :   |   | | :                  :| :| |   | |  | |::    :    |
Db         408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy         155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
                |   |    ||       :| ||         |  |: : ||    :  :|:
Db         453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC 511

Qy         201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                     : |:  : | : |  |    :||    |
```

```
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


RESULT 15
US-11-052-554A-273
; Sequence 273, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 273
;   LENGTH: 1255
;   TYPE: PRT
;   ORGANISM: SARS coronavirus GD01
US-11-052-554A-273

  Query Match             5.9%;  Score 86;  DB 7;  Length 1255;
  Best Local Similarity   21.4%;  Pred. No. 9.6;
  Matches   46;  Conservative   27;  Mismatches   92;  Indels   50;  Gaps   11;

Qy         42 SAKYNPS-ISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPG-----ID 95
              |  || :  | |: :    | :|  :     |:    |  : | ||  ::|||      |
Db        353 SVLYNSTFFSTFKCYGVSATKLN--DLCFSNVYA---DSFVVKGDDVRQIAPGQTGVIAD 407

Qy         96 FQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEY-YKHFALSRKDAME 154
              :   |   | | :                 :| :| |   | |  | |::    :    |
Db        408 YNYKLPDDFMGCV--------------LAWNTRNIDATSTGNYNYKYRYLRHGKLRPFE 452

Qy        155 -DQQYVVLKNDGITFMSLMVNTCY-------DITAEGVSFVPYACAGIGADLIT------ 200
               |   |    ||       :| ||         |  |: : ||    :  :|:
Db        453 RDISNVPFSPDGKPCTPPALN-CYWPLNDYGFYTTTGIGYQPYRVVVLSYELLNAPATV 511

Qy        201 -------IFKDLNLKFAYQGKIGISYPITPEVSAF 228
                    : |:  : | :|  |    :||    |
Db        512 GPKLSTDLIKNQCVNFNFNGLTGTGV-LTPSSKRF 545


Search completed: January 24, 2006, 12:04:18
Job time : 31.1111 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:36:23 ; Search time 43.9216 Seconds
(without alignments)
613.382 Million cell updates/sec

Title: US-10-680-349-42
Perfect score: 1462
Sequence: 1 MNYKKILVRSALISLMSILP.....ASVTLDVGYFGGEIGMRFTF 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	629.5	43.1	276	2	JE0218	28k surface antige
2	620	42.4	286	2	JE0219	28k surface antige
3	603	41.2	280	2	JE0217	28k surface antige
4	583	39.9	278	2	JE0216	28k surface antige
5	565	38.6	284	2	I40882	major antigenic pr
6	515.5	35.3	287	2	JE0220	28k surface antige
7	386.5	26.4	282	2	I39648	major surface prot
8	227.5	15.6	133	2	JE0221	28k surface antige
9	119.5	8.2	264	2	I54668	heat resistant agg
10	107.5	7.4	239	2	AH0541	probable outer mem
11	104.5	7.1	738	2	S01892	hemolysin A precur
12	102	7.0	1176	2	A27826	DNA-directed RNA p
13	101.5	6.9	415	2	B70448	3-oxoacyl-[acyl-ca
14	100	6.8	1582	2	AC1153	adhesin homolog 1m
15	98	6.7	3890	2	C89921	hypothetical prote
16	97.5	6.7	274	2	AC3295	heat resistant agg
17	97.5	6.7	1174	2	S28976	DNA-directed RNA p
18	97.5	6.7	1271	2	A45555	glutamate rich pro
19	97.5	6.7	1310	2	AD1380	glycosidase homolo
20	95	6.5	1004	2	B25039	outer cell wall pr
21	95	6.5	1578	2	AD1512	peptidoglycan boun
22	94.5	6.5	474	2	F97264	6-Phospho-Beta-D-G
23	93.5	6.4	658	2	E95111	endo-beta-N-acetyl
24	93	6.4	180	2	F71639	hypothetical prote
25	93	6.4	608	2	H64473	hypothetical prote
26	93	6.4	745	2	T51370	hypothetical prote
27	92.5	6.3	219	2	AF2658	conserved hypothet
28	92.5	6.3	219	2	C97440	hypothetical prote
29	92.5	6.3	401	2	F86754	prophage pi2 prote
30	92.5	6.3	585	2	C69336	probable electron
31	91.5	6.3	721	2	C97980	endo-beta-N-acetyl
32	91.5	6.3	1224	2	A25884	DNA-directed RNA p
33	91.5	6.3	1635	2	AI0452	hemolysin [importe
34	91	6.2	282	2	S16617	opacity protein op
35	91	6.2	588	2	AG0517	penicillin-binding
36	91	6.2	1348	2	B23496	TyB protein - yeas
37	91	6.2	1598	2	S69967	TyB protein - yeas
38	91	6.2	1770	2	S69953	TyB protein - yeas
39	91	6.2	1770	2	S58651	TyB protein - yeas
40	91	6.2	1770	2	S70233	TyB protein - yeas
41	91	6.2	1770	2	S69966	TyB protein - yeas
42	91	6.2	1770	2	S70230	TyB protein - yeas
43	91	6.2	1770	2	S69950	TyB protein - yeas
44	91	6.2	1771	2	S53592	TyB protein - yeas
45	91	6.2	1810	2	S69973	TyB protein - yeas

ALIGNMENTS

RESULT 1
JE0218
28k surface antigen 5 - Ehrlichia chaffensis
N;Alternate names: MAP1
C;Species: Ehrlichia chaffensis
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C;Accession: JE0218
R;Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A;Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe E
A;Reference number: JE0216; MUID:98321180; PMID:9647746
A;Accession: JE0218
A;Molecule type: DNA
A;Residues: 1-276 <RED>
A;Cross-references: UNIPROT:Q93DD2; UNIPROT:Q9ACI9; UNIPARC:UPI000003478F; GB:AF062761

Query Match 43.1%; Score 629.5; DB 2; Length 276;
Best Local Similarity 47.9%; Pred. No. 5.4e-44;
Matches 137; Conservative 41; Mismatches 87; Indels 21; Gaps 6;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                  : |  |||||:: |       :  | ||       ::||   ||:|:|||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGPAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||     | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
               |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIG 275
              |::| ||||:|  || | |     | :    | | ||| :|| |:|
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIEMG 276
```

RESULT 2
JE0219
28k surface antigen 2 - Ehrlichia chaffensis
N;Alternate names: MAP1
C;Species: Ehrlichia chaffensis
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C;Accession: JE0219
R;Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A;Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe E
A;Reference number: JE0216; MUID:98321180; PMID:9647746

```
A;Accession: JE0219
A;Molecule type: DNA
A;Residues: 1-286 <RED>
A;Cross-references: UNIPROT:O52105; UNIPARC:UPI000003478C; GB:AF062761

  Query Match             42.4%;  Score 620;  DB 2;  Length 286;
  Best Local Similarity   47.3%;  Pred. No. 3.4e-43;
  Matches  139;  Conservative   42;  Mismatches   89;  Indels   24;  Gaps    7;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEG-FYISAKYNPSISHFRKFSAEE 59
              || :|  : :||  ||| ||  | :|||    :||  | |||| || || |||  |||:|
Db          1 MNCEKFFITTALTLLMSFLPGISLSDPV---QDDNISGNFYISGKYMPSASHFGVFSAKE 57

Qy         60 TPINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMD 113
                : |  |||:::| |        |  | ||       ::||| |||:|:||||||
Db         58 -----ERNTTVGVFGIEQDWDRCVISRTTLSDIFTVPNYSFKYENNLFSGFAGAIGYSMD 112

Qy        114 GPRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQ------QYVVLKNDGIT 167
              ||||||| :|: |: ||  ||      | :::|||      | |         | | |:|:
Db        113 GPRIELEVSYEAFDVKNQGNN--YKNEAHRYYALSHLLGTETQIDGAGSASVFLINEGLL 170

Qy        168 FMSLMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSA 227
                | |:| |||: :||: | || ||||| ||:::|: :| | :||||:|:||||:|| |
Db        171 DKSFMLNACYDVISEGIPFSPYICAGIGIDLVSMFEAINPKISYQGKLGLSYPISPEASV 230

Qy        228 FIGGYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              ||||::| ||||:|  || : |      | :    | ||||| ||| |:| || |
Db        231 FIGGHFHKVIGNEFRDIPTMIPSESALAGKGNYPAIVTLDVFYFGIELGGRFNF 284


RESULT 3
JE0217
28k surface antigen 4 - Ehrlichia chaffensis
N;Alternate names: MAP1
C;Species: Ehrlichia chaffensis
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C;Accession: JE0217
R;Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A;Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A;Reference number: JE0216; MUID:98321180; PMID:9647746
A;Accession: JE0217
A;Molecule type: DNA
A;Residues: 1-280 <RED>
A;Cross-references: UNIPROT:O52107; UNIPARC:UPI000003478E; GB:AF062761

  Query Match             41.2%;  Score 603;  DB 2;  Length 280;
  Best Local Similarity   45.1%;  Pred. No. 8e-42;
  Matches  133;  Conservative   54;  Mismatches   78;  Indels   30;  Gaps   10;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEG-FYISAKYNPSISHFRKFSAEE 59
              || ||  : : |:|||| ||  ||:| |    |||  | |||| || ||:|||  |||:|
Db          1 MNCKKFFITTTLVSLMSFLPGISFSDAV---QNDNVGGNFYISGKYVPSVSHFGVFSAKQ 57

Qy         60 TPINGTNSLTKKVFGLKKDGD---ITKKD-DFTRVAPGIDF--QNNLISGFSGSIGYSMD 113
                : |  |||||:| |   |:|   : |   |   |  :||   ||:|::|| |:
Db         58 -----ERNTTIGVFGLKQDWDGSTISKNSPENTFNVPNYSFKYENNPFLGFAGAVGYLMN 112

Qy        114 GPRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSR----KDAMEDQQYVVLKNDGITFM 169
              ||||||| :|: |: ||  ||  ::   :|::||:     | :    ::| |||:|:  :
Db        113 GPRIELEMSYETFDVKNQGNNYKNDA--HKYYALTHNSGGKLSNAGDKFVFLKNEGLLDI 170

Qy        170 SLMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFI 229
              |||:| |||: :||: | || |||:| |||::|: :| | :||||:|:|| |:|| | |:
Db        171 SLMLNACYDVISEGIPFSPYICAGVGTDLISMFEAINPKISYQGKLGLSYSISPEASVFV 230

Qy        230 GGYYHGVIGNKFEKIPVITPVVLNDAPQTTSAS----VTLDVGYFGGEIGMRFTF 280
              ||::| ||||:|  || : |      :  | : :    ||| | :|| |:| || |
Db        231 GGHFHKVIGNEFRDIPAMIP-----STSTLTGNHFTIVTLSVCHFGVELGGRFNF 280


RESULT 4
JE0216
```

```
28k surface antigen 3 - Ehrlichia chaffensis
N;Alternate names: MAP1
C;Species: Ehrlichia chaffensis
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C;Accession: JE0216
R;Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A;Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe E
A;Reference number: JE0216; MUID:98321180; PMID:9647746
A;Accession: JE0216
A;Molecule type: DNA
A;Residues: 1-278 <RED>
A;Cross-references: UNIPROT:O52106; UNIPARC:UPI000003478D; GB:AF062761

  Query Match             39.9%;  Score 583;  DB 2;  Length 278;
  Best Local Similarity   43.6%;  Pred. No. 3.4e-40;
  Matches  126;  Conservative   49;  Mismatches   94;  Indels   20;  Gaps    7;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEG-FYISAKYNPSISHFRKFSAEE 59
              || ||  : :||:|||| ||  ||:|||      ||  | ||:| || || |||  |||:|
Db          1 MNCKKFFITTALVSLMSFLPGISFSDPV---QGDNISGNFYVSGKYMPSASHFGMFSAKE 57

Qy         60 TPINGTNSLTKKVFGLKKD-----GDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                    : |  ::|||:|          : |      ::||   ||:|:||||| |
Db         58 -----EKNPTVALYGLKQDWEGISSSSHNDNHFNNKGYSFKYENNPFLGFAGAIGYSMGG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKD---AMEDQQYVVLKNDGITFMSL 171
              ||:| | :|: |: ||  ||  ::   ::: || ::|      :  :||:||::|:  :|
Db        113 PRVEFEVSYETFDVKNQGNNYKNDA--HRYCALGQQDNSGIPKTSKYVLLKSEGLLDISF 170

Qy        172 MVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGG 231
              |:| ||||  | :   || |||:| |||::|:  | | :||||:|:|| | || | ||||
Db        171 MLNACYDIINESIPLSPYICAGVGTDLISMFEATNPKISYQGKLGLSYSINPEASVFIGG 230

Qy        232 YYHGVIGNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ::| ||||:|  || :   | : |     | ||| | :|| |:| || |
Db        231 HFHKVIGNEFRDIPTLKAFVTSSATPDL-AIVTLSVCHFGIELGGRFNF 278


RESULT 5
I40882
major antigenic protein - heartwater rickettsia
C;Species: Cowdria ruminantium (heartwater rickettsia)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40882; S42827
R;van Vliet, A.H.; Jongejan, F.; van Kleef, M.; van der Zeijst, B.A.
Infect. Immun. 62, 1451-1456, 1994
A;Title: Molecular cloning, sequence analysis, and expression of the gene encoding the in
A;Reference number: I40882; MUID:94178956; PMID:8132352
A;Accession: I40882
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-284 <RES>
A;Cross-references: UNIPROT:Q46327; UNIPARC:UPI00000B9A44; EMBL:X74250; NID:g454266; PIDN
C;Genetics:
A;Gene: map1

  Query Match             38.6%;  Score 565;  DB 2;  Length 284;
  Best Local Similarity   43.1%;  Pred. No. 1e-38;
  Matches  129;  Conservative   39;  Mismatches   97;  Indels   34;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFAD-------PVGSRTNDNKEGFYISAKYNPSISHFR 53
              || ||| : | ||||:| ||  ||:|       ||||             |||||  |: |||
Db          1 MNCKKIFITSTLISLVSFLPGVSFSDVIQEENNPVGS--------VYISAKYMPTASHFG 52

Qy         54 KFSAEETPINGTNSLTKKVFGLKKDGDITKKDD------FTRVAPGIDFQNNLISGFSGS 107
              | | :|   :     || |||||||  |  |          ||        ::||   ||:|:
Db         53 KMSIKEDSRD-----TKAVFGLKKDWDGVKTPSGNTNSIFTEKDYSFKYENNPFLGFAGA 107

Qy        108 IGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGIT 167
              :||||:||||| | :|: |: :|    |  ::    |      :           |::||: :|
Db        108 VGYSMNGPRIEFEVSYETFDVRNPGGNYKNDAHMYCALDTASSSTAGATTSVMVKNENLT 167
```

```
Qy        168 FMSLMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSA 227
              :|||:| ||||  :|:   || ||||| ||:::    | | :||||:|||| | || |
Db        168 DISLMLNACYDIMLDGMPVSPYVCAGIGTDLVSVINATNPKLSYQGKLGISYSINPEASI 227

Qy        228 PIGGYYHGVIGNKFEKIPVITPVVLNDAPQTTS------ASVTLDVGYFGGEIGMRFTF 280
              ||||::| ||||:|: |   |  |   :  :|      ||  ||| :|| ||| || | |
Db        228 PIGGHFHRVIGNEFKDI--ATSKVFTSSGNASSAVSPGFASAILDVCHFGIEIGGRFVF 284


RESULT 6
JE0220
28k surface antigen 1 - Ehrlichia canis
C;Species: Ehrlichia canis
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C;Accession: JE0220
R;Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A;Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A;Reference number: JE0216; MUID:98321180; PMID:9647746
A;Accession: JE0220
A;Molecule type: DNA
A;Residues: 1-287 <RED>
A;Cross-references: UNIPROT:Q9ZGJ0; UNIPARC:UPI0000034790; GB:AF062762; NID:g3327964; PI

  Query Match            35.3%;  Score 515.5;  DB 2;  Length 287;
  Best Local Similarity  39.3%;  Pred. No. 1.2e-34;
  Matches  120;  Conservative   47;  Mismatches   95;  Indels   43;  Gaps    7;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              | |||    :||: | |   :  | |  : |  |   |||| || |: |||  |||:|
Db          1 MKYKKTFTVTALVLLTSFTHFIPFYSPARASTIHN---FYISGKYMPTASHFGIFSAKE- 56

Qy         61 PINGTNSLTKKVFGL--KKDGDITKKDDFTRVAP----GIDFQNNLISGFSGSIGYSMDG 114
                 | || : ||  :   :|  :|  :          ::||   ||:|:||||:
Db         57 ----EQSFTKVLVGLDQRLSHNIINNNDTAKSLKVQNYSFKYKNNPFLGFAGAIGYSIGN 112

Qy        115 PRIELEAAYQQFNPKNTDNN------------------DTDNGEYYKHFALSRKDAMED 155
              ||||| ::: |: ||  ||                  | ::|::|          :
Db        113 SRIELEVSHEIFDTKNPGNNYLNDSHKYCALSHGSHICSDGNSGDWY---------TAKT 163

Qy        156 QQYVVLKNDGITFMSLMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKI 215
              ::|:|||:|:  :| |:| ||||| | : | || ||||| |||::|:    | :||||:
Db        164 DKFVLLKNEGLLDVSFMLNACYDITTEKMPFSPYICAGIGTDLISMPETTQNKISYQGKL 223

Qy        216 GISYPITPEVSAFIGGYYHGVIGNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIG 275
              |::| |   || | ||::| ||||:|: || : |   |  |  ||:|||| :|| |||
Db        224 GLNYTINSRVSVFAGGHFHKVIGNEFKGIPTLLPDGSNIKVQ-QSATVTLDVCHFGLEIG 282

Qy        276 MRFTF 280
              || |
Db        283 SRFFF 287


RESULT 7
I39648
major surface protein 4 - Anaplasma marginale
C;Species: Anaplasma marginale
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C;Accession: I39648
R;Oberle, S.M.; Barbet, A.F.
Gene 136, 291-294, 1993
A;Title: Derivation of the complete msp4 gene sequence of Anaplasma marginale without cl
A;Reference number: I39648; MUID:94124017; PMID:8294020
A;Accession: I39648
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-282 <RES>
A;Cross-references: UNIPROT:Q07408; UNIPARC:UPI0000000B7F; GB:L01987; NID:g142174; PIDN:
C;Genetics:
A;Gene: msp4

  Query Match            26.4%;  Score 386.5;  DB 2;  Length 282;
  Best Local Similarity  32.2%;  Pred. No. 3.9e-24;
  Matches   96;  Conservative   54;  Mismatches  113;  Indels   35;  Gaps    9;

Qy          1 MNYKKIL---VRSALISLMSILPYQS-FADP----VGSRTNDNKEGFYISAKYN---PSI 49
              |||:::    : :| :   |:|   :  | |    | |         ||: | |:   ||:
Db          1 MNYRELFTGGLSAATVCACSLLVSGAVVASPMSHEVASEGGVMGGSFYVGAAYSPAFPSV 60

Qy         50 SHF--RKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGS 107
              : |  |: | | : : | :     :       |::  :|::      |  |||: | |:
Db         61 TSFDMRESSKETSYVRGYDKSIATI-------DVSVPANFSKSGYTFAFSKNLITSFDGA 113

Qy        108 IGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGEYYKH-----FALSRKDAMEDQQYVVLK 162
              :|||: | |:||||:|::|           :|:| |       |::|   : :  | |:|
Db        114 VGYSLGGARVELEASYRRF-------ATLADGQYAKSGAESLAAITRDANITETNYFVVK 166

Qy        163 NDGITFMSLMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPIT 222
              | ||  |:|:| |||:    :   || ||||||  : | | :  | ||:||:|||  |
Db        167 IDEITNTSVMLNGCYDVLHTDLPVSPYVCAGIGASFVDISKQVTTKLAYRGKVGISYQFT 226

Qy        223 PEVSAFIGGYYHGVIGNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||:|   ||:|||:    :: ||    | :       |||   :  :|  :| || |
Db        227 PEISLVAGGFYHGLFDESYKDIPAHNSVKFSG---EAKASVKAHIADYGFNLGARFLF 281


RESULT 8
JE0221
28k surface antigen 2 - Ehrlichia canis
C;Species: Ehrlichia canis
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C;Accession: JE0221
R;Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A;Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A;Reference number: JE0216; MUID:98321180; PMID:9647746
A;Accession: JE0221
A;Molecule type: DNA
A;Residues: 1-133 <RED>
A;Cross-references: UNIPROT:O85360; UNIPARC:UPI0000034791; GB:AF062762; NID:g3327964; PII

  Query Match            15.6%;  Score 227.5;  DB 2;  Length 133;
  Best Local Similarity  41.4%;  Pred. No. 1.4e-11;
  Matches   58;  Conservative   20;  Mismatches   49;  Indels   13;  Gaps    4;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              || ||:   ||||| :  || |:::||    |     |||| || ||: ||  ||||
Db          1 MNCKKVFTISALISSIYFLPNVSYSNPVYG--NSMYGNFYISGKYMPSVPHFGIFSAEE- 57

Qy         61 PINGTNSLTKKVFGLKKD--GDI----TKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                    |  |:|||::  ||     :  |:||       : :|   ||: :||||:
Db         58 ----EKKKTTVVYGLKENWAGDAISSQSPDDNFTIRNYSFKYASNKFLGFAVAIGYSIGS 113

Qy        115 PRIELEAAYQQFNPKNTDNN 134
              ||||:| :|: |: ||  ||
Db        114 PRIEVEMSYEAFDVKNQGNN 133


RESULT 9
I54668
heat resistant agglutinin 1 - Escherichia coli
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C;Accession: I54668; I69133; I69134
R;Lutwyche, P.; Rupps, R.; Cavanagh, J.; Warren, R.A.; Brooks, D.E.
Infect. Immun. 62, 5020-5026, 1994
A;Title: Cloning, sequencing, and viscometric adhesion analysis of heat-resistant aggluti
A;Reference number: I54668; MUID:95012721; PMID:7927783
A;Accession: I54668
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-264 <RES>
A;Cross-references: UNIPROT:Q46678; UNIPARC:UPI00000BE5DA; EMBL:U07174; NID:g463910; PIDN
A;Accession: I69133
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
```

```
A;Residues: 15-264 <RE2>
A;Cross-references: UNIPARC:UPI00000B73D0; EMBL:U07174; NID:g463910; PIDN:AAC13752.1; PI
A;Accession: I69134
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 18-264 <RE3>
A;Cross-references: UNIPARC:UPI00000B8B4C; EMBL:U07174; NID:g463910; PIDN:AAC13753.1; PI

  Query Match             8.2%;  Score 119.5;  DB 2;  Length 264;
  Best Local Similarity   23.0%;  Pred. No. 0.023;
  Matches   59;  Conservative   41;  Mismatches   66;  Indels   91;  Gaps   14;

Qy          5 KILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEETPING 64
              |::  ||| ::  :   |: ||       ::| |||:: |   |:       |::  ::|
Db         20 KVIAVSAL-AMAGMFSTQALAD-------ESKTGFYVTGKAGASVMSL----ADQRFLSG 67

Qy         65 TNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMD-------GPRI 117
                | |    | ||           | |      : ||| |    |            |
Db         68 NGEETSKY----KGGD-----------GHD------TVFSGGIAAGYDFYPQFSIPVRT 105

Qy        118 ELE-AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTC 176
              |||  |  : : |   : |: :| |::          :|     |||: ::  :||:|
Db        106 ELEFYARGKADSKYNVDKDSWSGGYWR----------DD-----LKNE-VSVNTLMLNAY 149

Qy        177 YDITAEGVSFVPYACAGIGADLITIFKDLNLK--------------------------F 209
              ||   :  :| |:  ||||       |::: |                          |
Db        150 YDFRNDS-AFTPWVSAGIG-----YAKEIHQKTTGISTWDYGYGSSGRESLSRSGSADNF 203

Qy        210 AYQGKIGISYPITPEVS 226
              |:    |: | :||:::
Db        204 AWSLGAGVRYDVTPDIA 220


RESULT 10
AH0541
probable outer membrane adhesin STY0351 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0541
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0541
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-239 <PAR>
A;Cross-references: UNIPARC:UPI000005A293; GB:AL513382; PIDN:CAD08776.1; PID:g16501592;
C;Genetics:
A;Gene: STY0351

  Query Match             7.4%;  Score 107.5;  DB 2;  Length 239;
  Best Local Similarity   22.6%;  Pred. No. 0.19;
  Matches   59;  Conservative   38;  Mismatches   85;  Indels   79;  Gaps   15;

Qy         30 SRTNDNKEGFYISAKYNPSISHF----RKFSAEETPINGTNSL---TKKVFGLKKDGDIT 82
              | |   ||| ||: |   |: :       || ||  :||  :|   || |||    | :
Db         17 SATASAKEGIYITGKAGTSVVNVYGINSTFSQEEI-VNGHATLPDRTKGVFG----GGV- 70

Qy         83 KKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGP-----RIELEAAYQQFNPKNTDNNDTD 137
                                :|||     |     |:||:  ::           :||
Db         71 ------------------------AIGYDFYDPFQLPVRLELDTTFR---------GETD 97

Qy        138 NGEYYKHFALSRKD--AMEDQQYVVLKNDGITFMSLMVNTCYDITAEGVSFVPYACAGIG 195
                    :|  |     :: :||  :   : :||  ||      :| ||  ||:|
Db         98 --------AKGGQDIIAFGQPVHINVKNQ-VRMTTYIVNGYYDF-HNSTAFTPYISAGVG 147

Qy        196 ADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGV-IGNKFEKIPVITPVVLND 254
                  : :  :          | : : |   :||    :  |  || |:     :|  :: |
Db        148 LAHVKLSNN---------TIPVGFGINETLSASKNNFAWGAGIGAKY----AVTDNIMID 194

Qy        255 A--PQTTSASVTLDVGYFGGE 273
                  |       :  |::   :: |:
Db        195 ASYKYINAGKVSISKNHYAGD 215


RESULT 11
S01892
hemolysin A precursor - Vibrio cholerae (strain El Tor 017)
N;Alternate names: El Tor hemolysin
C;Species: Vibrio cholerae
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 10-Dec-1999
C;Accession: S01892
R;Alm, R.A.; Stroeher, U.H.; Manning, P.A.
Mol. Microbiol. 2, 481-488, 1988
A;Title: Extracellular proteins of Vibrio cholerae: nucleotide sequence of the structural
in the non-haemolytic classical strain 569B.
A;Reference number: S01892; MUID:89013889; PMID:3050359
A;Accession: S01892
A;Molecule type: DNA
A;Residues: 1-738 <ALM>
A;Cross-references: UNIPARC:UPI000017918E; EMBL:Y00557
C;Genetics:
A;Gene: hlyA
C;Superfamily: Vibrio hemolysin A
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-738/Product: hemolysin A #status predicted <MAT>

  Query Match             7.1%;  Score 104.5;  DB 2;  Length 738;
  Best Local Similarity   21.5%;  Pred. No. 1.5;
  Matches   63;  Conservative   37;  Mismatches   98;  Indels   95;  Gaps   14;

Qy         28 VGSRTNDNK------------EGFYIS---------AKYNPSISHFRKFSAEETP----- 61
              ||| | | |              | :::         | |    ::||::| :
Db        220 VGSATPDAKIVRISLDDDSTGAGIHLNDQLGYRQFGASYTTLDAYFREWSTDAIAQDYRF 279

Qy         62 -INGTNSLTK--KVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGF----SGSIGYSMDG 114
               | :|:  :  | | :   :| :|           |:   :|||    :| :  | ||
Db        280 VFNASNNKAQILKTFPV---DNINEK----------FERKEVSGFELGVTGGVEVSGDG 325

Qy        115 PRIELEAAYQQFNPKNTDNNDTD-----NGEYYKHFALSRKDAMEDQQYVVLKNDGITFM 169
              |: :|||  : :  :    |  |     | :  |  ||     : | :         :|
Db        326 PKAKLEARRKLYQSRWLTYNTQDYRIERNAKNAKRLALHELSTIRDSRIATQ-----SFD 380

Qy        170 SLMVNTCYDITAEGVSFVPYA------------CAGIGADLITIFKDLNLKFAYQGKIG 216
                  |  | :    :| : ||                  |:    |   :|::  | |
Db        381 RCFVGEYYPVDVNRISALTYASFVPKMDVIYKASATETGSTDFIIDSSVNIRPIYNGAYK 440

Qy        217 ISYPITPEVSAFIGGYYHGVIGNKFEKIPVITPVVLNDAPQ---TTSASVTLD 266
              | :   |     |||     ||       | |:   | ||| |:|
Db        441 HYYVVGAHQS------YHG-----FE----------DTPRRRITKSASFTVD 471


RESULT 12
A27826
DNA-directed RNA polymerase (EC 2.7.7.6) II second largest chain - fruit fly (Drosophila
C;Species: Drosophila melanogaster
C;Date: 19-May-1989 #sequence_revision 23-Nov-1991 #text_change 09-Jul-2004
C;Accession: A27826; PQ0154
R;Falkenburg, D.; Dworniczak, B.; Faust, D.M.; Bautz, E.K.F.
J. Mol. Biol. 195, 929-937, 1987
A;Title: RNA polymerase II of Drosophila. Relation of its 140,000 Mr subunit to the beta
A;Reference number: A27826; MUID:88011299; PMID:3116266
A;Accession: A27826
A;Molecule type: DNA
A;Residues: 54-1176 <FAL>
A;Cross-references: UNIPROT:P08266; UNIPARC:UPI000016BD32; GB:X05709; GB:M29646; NID:g551
R;Sitzler, S.; Oldenburg, I.; Petersen, G.; Bautz, E.K.F.
Gene 100, 155-162, 1991
A;Title: Analysis of the promoter region of the housekeeping gene DmRP140 by sequence com
A;Reference number: PQ0154; MUID:91276237; PMID:1905256
A;Accession: PQ0154
```

```
A;Molecule type: mRNA
A;Residues: 1-69 <SIT>
A;Cross-references: UNIPARC:UPI000016BB8E; GB:M62972; NID:g157263; PIDN:AAA28476.1; PID:
A;Experimental source: embryo
C;Genetics:
A;Gene: DmRP140
A;Cross-references: FlyBase:FBgn0003276
A;Introns: 8/1
C;Superfamily: DNA-directed RNA polymerase 132K polypeptide
C;Keywords: DNA binding; nucleotidyltransferase; nucleus; transcription; zinc finger

  Query Match             7.0%;  Score 102;  DB 2;  Length 1176;
  Best Local Similarity   19.9%;  Pred. No. 4.6;
  Matches   53;  Conservative   39;  Mismatches   78;  Indels   96;  Gaps   12;

Qy         21 YQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEETPINGTNSLTKKVFGLKKDGD 80
              |:|: |    |  | :| |                 | |  ||       |: ::
Db        813 YRSYKDSENKRVGDQEENF-----------------EKPHRGTCQ------GM-RNAH 846

Qy         81 ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELEAAYQQFNPKNTDNNDTDNGE 140
               | ||   :|||||      :||    || ::               |:| |  |::
Db        847 YDKLDDDGIIAPGI-----RVSGDDVVIGKTIT------------LPENDDELDSNTKR 888

Qy        141 YYKH----FALSRKDAMEDQQYVVLKNDGITFMSLMV-----------------NTCY 177
              : |     | : :  : || : | ::|  |  : |                    ||
Db        889 FSKRDASTFLRNSETGIVDQVMLTLNSEGYKFCKIRVRSVRIPQIGDKFASRHGQKGTC- 947

Qy        178 DITAEGVSF----VPYACAGIGADLIT-------------IFKDLNLKF-AYQGKIGISY 219
              |: :    : : | |:  |:|                 : : |  |  : :|:|| :
Db        948 -----GIQYRQEDMAFTCEGLAPDIIINPHAIPSRMTIGHLIECLQGKLGSNKGEIGDAT 1002

Qy        220 PI-----TPEVSAFIGGYYHGVIGNK 240
              |        ::| |:  | : : ||:
Db       1003 PFNDAVNVQKISTFLQEYGYHLRGNE 1028


RESULT 13
B70448
3-oxoacyl-[acyl-carrier-protein] synthase (EC 2.3.1.41) II - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: B70448
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: B70448
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-415 <AQF>
A;Cross-references: UNIPROT:O67612; UNIPARC:UPI00000566C3; GB:AE000752; NID:g2984021; PI
A;Experimental source: strain VF5
C;Genetics:
A;Gene: fabF
C;Function:
A;Pathway: fatty acid biosynthesis
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I; 3-oxoacyl-[acyl-carrier-prot
C;Keywords: acyltransferase; fatty acid biosynthesis
F;23-410/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F;165/Active site: Cys #status predicted

  Query Match             6.9%;  Score 101.5;  DB 2;  Length 415;
  Best Local Similarity   21.7%;  Pred. No. 1.3;
  Matches   60;  Conservative   32;  Mismatches   73;  Indels  111;  Gaps   14;

Qy         62 INGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSG-----SIGYSMDGPR 116
              | |  ::|    |:||                 | ||: | ||      |       |
Db          7 ITGIGAVTPLGTGVKK-----------------FWENLVKGVSGVDYIKSFNPDEYGIP 48

Qy        117 IELEAAYQQFNPKN-TDNNDTDNGEYYKHFALSR-KDAMED------------------ 155
              ::: |  : |||::  |  :      :  ||::  |:|:||
Db         49 VKIAAEVKDFNPEDFMDKKEARKASRFVQFAIAAVKEALEDSGLLESEYDPYRVGVIIGT 108

Qy        156 ---------QQYVVLKNDGITFMSLMVNTCYDITAEGVS--FVPYACAGIGADLITIFKD 204
                      | ::||  |                |  ||  |:||  : : : |: |
Db        109 GIGGLKDIEDQTLILKEKG--------------ARRVSPFFIPYGISNMASGLVAI--- 150

Qy        205 LNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNKFEKI-------------PVITPV 250
                 |: ::|    :| :|   ||    | :   ||: |  |                |||:
Db        151 ---KWGFKGP---NYCVT---SACATGNH--AIGDAFRLIQKGDIDIAIAGGTEAAITPL 199

Qy        251 -----------VLNDAPQTTSASVTLD-VGYFGGE 273
                        || ||  |     |  |:  ||
Db        200 GVAGPASMKALSTRNDEPQKASRPFDRDRDGFVMGE 235


RESULT 14
AC1153
adhesin homolog lmo0627 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC1153
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1153
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1582 <GLA>
A;Cross-references: UNIPROT:Q8Y9A5; UNIPARC:UPI0000054C9E; GB:NC_003210; PIDN:CAC98705.1,
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0627

  Query Match             6.8%;  Score 100;  DB 2;  Length 1582;
  Best Local Similarity   21.4%;  Pred. No. 9.9;
  Matches   63;  Conservative   30;  Mismatches   84;  Indels  118;  Gaps   13;

Qy         39 FYISAKYNPSISHFRKFSAEETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQN 98
              ||:: :    |  |  |: | |                  || ||  |    : ||    |
Db        548 FYMAGE----IGSFSNFNGEST---------------LDGAITDVDSTPDMNPG----N 583

Qy         99 NLISGFSGSI-----GYSMDGPRIELEAAYQQFNPKNTDNNDTD------------NGE 140
              :|||   ||      |: : | :  | |       :: |: | |             |:
Db        584 DLISNVDGSTIKPIDGFQKEKPHVVKENAKSTATLQDEDDFDFDYVTFIKTPEVIPSEGK 643

Qy        141 YYKHFALSRKDAMED--------------QQYVVLKNDG-----ITFMSLMVN------- 174
               ::   ||  || :                : |  | :||      |:|   :|
Db        644 VFEKTRLSAADAEKTARNMIGNDIMGSALKDYTPLADDGRITSPKTYMVYEMNINPSGVE 703

Qy        175 -----TCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFI 229
                   :  |   :|:  : |  ||: :
Db        704 DTLNSSFTDTLPKGLKMLEYDIAGLNSS------------------------------ 731

Qy        230 GGYYHGVIGNKFEKIPVIT-----PVV-----LNDAPQTTSASVTL-DVG--YFG 271
               : :|   ||||  |  |       |:     |    |   |: || |||  |||
Db        732 NKHIYGFTTNKFEGTPKFTNADGQDVIVYQKGLYSEKQVCIANPTLNDVGVRYFG 786


RESULT 15
C89921
hypothetical protein ebhB [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: C89921
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguch
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; H
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
```

```
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89921
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3890 <KUR>
A;Cross-references: UNIPROT:Q99U53; UNIPARC:UPI00000CAB07; GB:BA000018; PID:g13701233; P
A;Experimental source: strain N315
C;Genetics:
A;Gene: ebhB

  Query Match             6.7%;  Score 98;  DB 2;  Length 3890;
  Best Local Similarity   22.3%;  Pred. No. 47;
  Matches   65;  Conservative   48;  Mismatches  114;  Indels   64;  Gaps   14;

Qy          3 YKKILVRSALISLMSILPYQ---SFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEE 59
              | |: | | : : ::|||     :  :|||           |:|||         :  :|
Db        892 YNKV-VASNMSNAVTILPDDIPPTINNPVG----------INAKY---------YRGDE 930

Qy         60 TPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIEL 119
               :| |  :: :  |:|     |    :|       | :|       ||:  :    |: :
Db        931 --VNFTMGVSDRHSGIKNTTITTLPSGWTSNLTKSDNKN-------GSLAIT---GRVSM 978

Qy        120 EAAYQQ---FNPKNTDN-NDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNT 175
               |:     |    ||| |:| |    || ::      ||   :|| |   |   ::||
Db        979 NQAFNSDITFKVSATDNVNNTTNDSQSKHVSIHVGKISEDAHPIVLGN---TEKVVVNP 1035

Qy        176 CYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHG 235
                    |  |            :|| | : |     :| :  : |:| : :  :  :|
Db       1036 TAVSNDEKQS-----------IITAFMNKNQNI--RGYLASTDPVTVDNNGNVTLHYRD 1081

Qy        236 VIGNKFEKIPVIT--PVVLND-----APQTTSASVTLDVGYFGGEIGMRFT 279
                 :   |:|  |||·::     | :| : ::      :  |:|   ||
Db       1082 GSSTTLDATNVMTYEPVVKSEYQTANAAKTATVTIAKGQSFNIGDIKQYFT 1132


Search completed: January 24, 2006, 11:47:47
Job time : 45.9216 secs
```

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:35:23 ; Search time 210.458 Seconds
(without alignments)
938.659 Million cell updates/sec

Title: US-10-680-349-42
Perfect score: 1462
Sequence: 1 MNYKKILVRSALISLMSILP.....ASVTLDVGYFGGEIGMRFTF 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1462	100.0	280	2	Q9F476_EHRCA	Q9f476 ehrlichia c
2	1452	99.3	280	2	Q84HU1_EHRCA	Q84hu1 ehrlichia c
3	1224.5	83.8	283	2	Q8G8D6_EHRCH	Q8g8d6 ehrlichia c
4	1217.5	83.3	283	2	Q8G8I2_EHRCH	Q8g8i2 ehrlichia c
5	1217.5	83.3	283	2	Q8GGU4_EHRCH	Q8ggu4 ehrlichia c
6	1202.5	82.3	283	2	O85358_EHRCH	O85358 ehrlichia c
7	1124	76.9	282	2	Q9R443_EHRRU	Q9r443 ehrlichia r
8	1123	76.8	282	2	Q5W5Y1_EHRRU	Q5w5y1 ehrlichia r
9	1121	76.7	282	2	Q4L0C0_EHRRU	Q4l0c0 ehrlichia r
10	1121	76.7	304	2	Q5FGL4_EHRRG	Q5fgl4 ehrlichia r
11	1121	76.7	304	2	Q5FCD9_EHRRW	Q5fcd9 ehrlichia r
12	1120	76.6	282	2	Q5HA10_EHRRW	Q5ha10 ehrlichia r
13	1120	76.6	282	2	Q9WW41_EHRRU	Q9ww41 ehrlichia r
14	665.5	45.5	275	2	Q93DD4_EHRCH	Q93dd4 ehrlichia c
15	655	44.8	276	2	Q8GGU0_EHRCH	Q8ggu0 ehrlichia c
16	651	44.5	276	2	O85817_EHRCH	O85817 ehrlichia c
17	648	44.3	280	2	Q9ZGM9_EHRCH	Q9zgm9 ehrlichia c
18	647	44.3	276	2	Q93DD1_EHRCH	Q93dd1 ehrlichia c
19	645.5	44.2	281	2	Q93DD2_EHRCH	Q93dd2 ehrlichia c
20	644.5	44.1	281	2	Q9ACI9_EHRCH	Q9aci9 ehrlichia c
21	644	44.0	280	2	O85816_EHRCH	O85816 ehrlichia c
22	628	43.0	280	2	Q93DD3_EHRCH	Q93dd3 ehrlichia c
23	621	42.5	288	2	Q9ZGJ2_EHRCA	Q9zgj2 ehrlichia c
24	620	42.4	286	2	O52105_EHRCH	O52105 ehrlichia c
25	605	41.4	280	2	O52107_EHRCH	O52107 ehrlichia c
26	605	41.4	291	2	Q8G8P3_EHRCH	Q8g8p3 ehrlichia c
27	601	41.1	280	2	Q9ADV3_EHRCA	Q9adv3 ehrlichia c
28	600	41.0	280	2	Q9F473_EHRCA	Q9f473 ehrlichia c
29	599	41.0	291	2	Q8G921_EHRCH	Q8g921 ehrlichia c
30	598	40.9	291	2	Q8GGU2_EHRCH	Q8ggu2 ehrlichia c
31	592	40.5	285	2	Q9L6Y5_EHRCH	Q9l6y5 ehrlichia c
32	589	40.3	287	2	Q8GGU3_EHRCH	Q8ggu3 ehrlichia c
33	588	40.2	278	2	Q9R3J3_EHRCA	Q9r3j3 ehrlichia c
34	588	40.2	278	2	Q9R8A5_EHRCA	Q9r8a5 ehrlichia c
35	588	40.2	278	2	Q9R8A6_EHRCA	Q9r8a6 ehrlichia c
36	588	40.2	278	2	Q9R8A7_EHRCA	Q9r8a7 ehrlichia c
37	588	40.2	278	2	Q9R8A8_EHRCA	Q9r8a8 ehrlichia c
38	588	40.2	278	2	Q9R8A9_EHRCA	Q9r8a9 ehrlichia c
39	588	40.2	307	2	Q9ZGJ1_EHRCA	Q9zgj1 ehrlichia c
40	583.5	39.9	277	2	Q8G8W7_EHRCH	Q8g8w7 ehrlichia c
41	583.5	39.9	277	2	Q8GGU1_EHRCH	Q8ggu1 ehrlichia c
42	583	39.9	278	2	Q9F472_EHRCA	Q9f472 ehrlichia c
43	583	39.9	278	2	O52106_EHRCH	O52106 ehrlichia c
44	573.5	39.2	279	2	Q8G8Q1_EHRCH	Q8g8q1 ehrlichia c
45	573.5	39.2	282	2	Q4W4X0_EHRRU	Q4w4x0 ehrlichia r

ALIGNMENTS

RESULT 1
Q9F476_EHRCA
ID Q9F476_EHRCA PRELIMINARY; PRT; 280 AA.
AC Q9F476;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE P28-2 (Major outer membrane protein P30-10).
GN Name=p28-2; Synonyms=p30-10;
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=944;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Jake;
RX MEDLINE=99242757; PubMed=10225842;
RA McBride J.W., Yu, Xj, Walker D.H.;
RT "Molecular cloning of the gene for a conserved major immunoreactive
RT 28-kilodalton protein of Ehrlichia canis: a potential serodiagnostic
RT antigen.";
RL Clin. Diagn. Lab. Immunol. 6:392-399(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Jake;
RX MEDLINE=20432107; PubMed=10974556; DOI=10.1016/S0378-1119(00)00256-0;
RA McBride J.W., Yu X.J., Walker D.H.;
RT "A conserved, transcriptionally active p28 multigene locus of
RT Ehrlichia canis.";
RL Gene 254:245-252(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oklahoma;
RX MEDLINE=98371112; PubMed=9705412;
RA Ohashi N., Unver A., Zhi N., Rikihisa Y.;
RT "Cloning and characterization of multigenes encoding the
RT immunodominant 30-kilodalton major outer membrane proteins of
RT Ehrlichia canis and application of the recombinant protein for
RT serodiagnosis.";
RL J. Clin. Microbiol. 36:2671-2680(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oklahoma;
RX MEDLINE=21153566; PubMed=11254561;
RX DOI=10.1128/IAI.69.4.2083-2091.2001;
RA Ohashi N., Rikihisa Y., Unver A.;
RT "Analysis of transcriptionally active gene clusters of major outer
RT membrane protein multigene family in Ehrlichia canis and E.
RT chaffeensis.";
RL Infect. Immun. 69:2083-2091(2001).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Venezuela, Arizona, California, and New Mexico;
RX MEDLINE=22461952; PubMed=12574308; DOI=10.1128/JCM.41.2.886-888.2003;

```
RA   Felek S., Greene R., Rikihisa Y.;
RT   "Transcriptional Analysis of p30 Major Outer Membrane Protein Genes of
RT   Ehrlichia canis in Naturally Infected Ticks and Sequence Analysis of
RT   p30-10 of E. canis from Diverse Geographic Regions.";
RL   J. Clin. Microbiol. 41:886-888(2003).
DR   EMBL; AF082744; AAG14357.1; -; Genomic_DNA.
DR   EMBL; AF078553; AAK28696.1; -; Genomic_DNA.
DR   EMBL; AF324792; AAK31312.1; -; Genomic_DNA.
DR   EMBL; AF528512; AAO41109.1; -; Genomic_DNA.
DR   EMBL; AF528515; AAO41112.1; -; Genomic_DNA.
DR   EMBL; AF528514; AAO41111.1; -; Genomic_DNA.
DR   EMBL; AF528511; AAO41108.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   280 AA;  30962 MW;  D1BB28B5AF0E6CA3 CRC64;

  Query Match             100.0%;  Score 1462;  DB 2;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-108;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 2
Q84HU1_EHRCA
ID   Q84HU1_EHRCA PRELIMINARY;      PRT;   280 AA.
AC   Q84HU1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major outer membrane protein.
GN   Name=p30-10;
OS   Ehrlichia canis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=944;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Hawaii;
RX   MEDLINE=22461952; PubMed=12574308; DOI=10.1128/JCM.41.2.886-888.2003;
RA   Felek S., Greene R., Rikihisa Y.;
RT   "Transcriptional Analysis of p30 Major Outer Membrane Protein Genes of
RT   Ehrlichia canis in Naturally Infected Ticks and Sequence Analysis of
RT   p30-10 of E. canis from Diverse Geographic Regions.";
RL   J. Clin. Microbiol. 41:886-888(2003).
DR   EMBL; AF528513; AAO41110.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   280 AA;  30990 MW;  D1BB28B14F5BDCA2 CRC64;

  Query Match             99.3%;  Score 1452;  DB 2;  Length 280;
  Best Local Similarity   99.3%;  Pred. No. 3.3e-107;
  Matches  278;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEEQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||| |||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLKDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 3
Q8G8D6_EHRCH
ID   Q8G8D6_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   Q8G8D6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   28kDa outer membrane protein gene 14.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479835; AAO12939.1; -; Genomic_DNA.
DR   EMBL; AF479836; AAO12944.1; -; Genomic_DNA.
DR   EMBL; AF479838; AAO12954.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31090 MW;  D5825DD4DC51C425 CRC64;

  Query Match             83.8%;  Score 1224.5;  DB 2;  Length 283;
  Best Local Similarity   81.6%;  Pred. No. 4.1e-89;
  Matches  231;  Conservative   24;  Mismatches   25;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| |||||||||||||||||||| |     ||:||||||| ||||||||||||||
Db          1 MNYKKIFVRSALISLMSILPYQSFADPVTSNDTGINDSKEGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
              |||||||| :|||||||||||||| :  :| |  | ::|||||||||||||:||||||
Db         61 EETPINGTTALTKKVFGLKKDGDIAQHGNFDRTDPALEFQNNLISGFSGSGIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||||||| :||||||:|||||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGIGADLINVFKDLNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||| |||||||:| |||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGIRFTF 283


RESULT 4
```

```
Q8G8I2_EHRCH
ID   Q8G8I2_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   Q8G8I2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   28kDa outer membrane protein gene 14.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479839; AAO12959.1; -; Genomic_DNA.
DR   EMBL; AF479840; AAO12965.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31069 MW;  AEA8B2833E02631F CRC64;

  Query Match             83.3%;  Score 1217.5;  DB 2;  Length 283;
  Best Local Similarity   80.9%;  Pred. No. 1.5e-88;
  Matches  229;  Conservative   24;  Mismatches   27;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| |||||||:||||||||||| |     ||:||||||| ||||||||||||||
Db          1 MNYKKIFVRSALISLVSILPYQSFADPVTSNDTGINDSKEGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              |||||||  ::||||||||||||| :  ||||  | ::|||||||||||||:||||||
Db         61 EETPINGNTAITKKVFGLKKDGDIAQHGDFTRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |:|||::|:|||:| ||| ||| |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDSNDTNSGDYYKYFGLSRDDAMADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||||||| :||| ||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGIGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||| |||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGIRFTF 283


RESULT 5
Q8GGU4_EHRCH
ID   Q8GGU4_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   Q8GGU4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   28kDa outer membrane protein gene 14.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479837; AAO12949.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31042 MW;  A2960DD4DC45942A CRC64;

  Query Match             83.3%;  Score 1217.5;  DB 2;  Length 283;
  Best Local Similarity   81.3%;  Pred. No. 1.5e-88;
  Matches  230;  Conservative   24;  Mismatches   26;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| |||||||||||||||||||||| |     ||:||||||| ||||||||||||||
Db          1 MNYKKIFVRSALISLMSILPYQSFADPVTSNDTGINDSKEGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              |||||||| :|||||||||||||| :  :| |  | ::|||||||||||||:||||||
Db         61 EETPINGTTALTKKVFGLKKDGDIAQHGNFDRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:  |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYVGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||||||| :||||||:|||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGIGADLINVFKDLNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | |||||||||| |||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGIRFTF 283


RESULT 6
O85358_EHRCH
ID   O85358_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   O85358; O52103;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   28kDa outer membrane protein gene 14 (Major outer membrane protein
DE   OMP-1B).
GN   Name=omp-1B;
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=98321180; PubMed=9647746; DOI=10.1006/bbrc.1998.8844;
RA   Reddy G.R., Sulsona C.R., Barbet A.F., Mahan S.M., Burridge M.J.,
RA   Alleman A.R.;
RT   "Molecular characterization of a 28 kDa surface antigen gene family of
RT   the tribe Ehrlichiae.";
RL   Biochem. Biophys. Res. Commun. 247:636-643(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Cheng C., Paddock C.D., Ganta R.R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Arkansas;
RX   MEDLINE=98084465; PubMed=9423849;
RA   Ohashi N., Zhi N., Zhang Y., Rikihisa Y.;
RT   "Immunodominant major outer membrane proteins of Ehrlichia chaffeensis
RT   are encoded by a polymorphic multigene family.";
RL   Infect. Immun. 66:132-139(1998).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Arkansas;
RX   MEDLINE=21153566; PubMed=11254561;
```

```
RX   DOI=10.1128/IAI.69.4.2083-2091.2001;
RA   Ohashi N., Rikihisa Y., Unver A.;
RT   "Analysis of transcriptionally active gene clusters of major outer
RT   membrane protein multigene family in Ehrlichia canis and E.
RT   chaffeensis.";
RL   Infect. Immun. 69:2083-2091(2001).
DR   EMBL; AF479833; AAO12929.1; -; Genomic_DNA.
DR   EMBL; U72291; AAC02936.1; -; Genomic_DNA.
DR   EMBL; AF479834; AAO12933.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31017 MW;  DCBCF652B771C95D CRC64;

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 2.3e-87;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::|||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| |||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 7
Q9R443_EHRRU
ID   Q9R443_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q9R443;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Major antigenic protein 1 like protein.
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99216274; PubMed=10198207; DOI=10.1006/bbrc.1999.0459;
RA   Sulsona C.R., Mahan S.M., Barbet A.F.;
RT   "The map1 gene of Cowdria ruminantium is a member of a multigene
RT   family containing both conserved and variable genes.";
RL   Biochem. Biophys. Res. Commun. 257:300-305(1999).
DR   EMBL; AF125279; AAD26353.1; -; Genomic_DNA.
DR   EMBL; AF125274; AAD26343.1; -; Genomic_DNA.
DR   EMBL; AF125277; AAD26349.1; -; Genomic_DNA.
DR   EMBL; AF125278; AAD26351.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   282 AA;  31040 MW;  0CB3C655775CBC90 CRC64;

  Query Match             76.9%;  Score 1124;  DB 2;  Length 282;
  Best Local Similarity   76.0%;  Pred. No. 4e-81;
  Matches  215;  Conservative   26;  Mismatches   38;  Indels    4;  Gaps    3;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRT--NDN-KEGFYISAKYNPSISHFRKFSA 57
              |||||||||||||||| |||||||:|| |    |:| ||||||||||||| |||||||
Db          1 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              |||: | :| |||||||||:| |||  ||||     : ||| ||||||||| |||||:
Db         61 EETPVYGKDSPTKKVFGLKKEGSITKYSDFTRTDISFEGQNNFISGFSGSIGYIMDGPRV 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |:|||||:|||||   |:||  :||||: ||| :|| |::|||| |:|:|| ||| | ||
Db        121 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAEAMADKKYVVLTNNGVTFSSLMFNACY 179

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:|||||||||||:|| |:|||||||||||||||||||:||||||||||||
Db        180 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 239

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||: |:||  || | ||||:|||||||| ||||||:|:||||
Db        240 GNKYNKVPVKLPVTLTDAPQSTSASVTLDAGYFGGELGVRFTF 282


RESULT 8
Q5W5Y1_EHRRU
ID   Q5W5Y1_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q5W5Y1;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   MAP1-1.
GN   Name=map1-1;
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Senegal;
RX   MEDLINE=22035368; PubMed=12039046; DOI=10.1016/S0378-1119(02)00408-0;
RA   Bekker C.P., Bell-Sakyi L., Paxton E.A., Martinez D., Bensaid A.,
RA   Jongejan F.;
RT   "Transcriptional analysis of the major antigenic protein 1 multigene
RT   family of Cowdria ruminantium.";
RL   Gene 285:193-201(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Senegal;
RX   PubMed=15995193; DOI=10.1128/JB.187.14.4782-4791.2005;
RA   Bekker C.P., Postigo M., Taoufik A., Bell-Sakyi L., Ferraz C.,
RA   Martinez D., Jongejan F.;
RT   "Transcription Analysis of the Major Antigenic Protein 1 Multigene
RT   Family of Three In Vitro-Cultured Ehrlichia ruminantium Isolates.";
RL   J. Bacteriol. 187:4782-4791(2005).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Senegal;
RA   Bekker C.P.J., Taoufik A., Jongejan F.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF319940; AAV54088.2; -; Genomic_DNA.
SQ   SEQUENCE   282 AA;  31097 MW;  3D619510E7393EBB CRC64;

  Query Match             76.8%;  Score 1123;  DB 2;  Length 282;
  Best Local Similarity   76.0%;  Pred. No. 4.9e-81;
  Matches  215;  Conservative   23;  Mismatches   41;  Indels    4;  Gaps    2;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDN---KEGFYISAKYNPSISHFRKFSA 57
              |||||||||||||||| |||||||:|| |    |   ||||||||||||| |||||||
Db          1 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNEXAKEGFYISAKYNPSIPHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              |||: | :| |||||||||:| |||  ||||     : ||| ||||||||| |||||:
Db         61 EETPVYGKDSPTKKVFGLKKDGSITKYSDFTRTDISFEGQNNFISGFSGSIGYIMDGPRV 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |:|||||:|||||   |:||  :||||: ||| : | |::|||| |:|:|| ||| | ||
Db        121 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAETMTDKKYVVLTNNGVTFSSLMFNACY 179

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
```

```
                ||||||| |:||||||||||||:|| |:||||||||||||||||||||:||||||||||||
Db        180 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 239

Qy        238 GNKPEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
                |||: |||| || | ||||:|||||||| ||||||:|:||||
Db        240 GNKYNKIPVKLPVTLTDAPQSTSASVTLDAGYFGGELGVRFTF 282


RESULT 9
Q4L0C0_EHRRU
ID   Q4L0C0_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q4L0C0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Map1-1.
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Gardel;
RA   Martinez D., Ferraz C., Demaille J., Bensaid A., Frutos R.;
RT   "Analysis of the major antigenic protein 1 multigene family of
RT   Ehrlichia ruminantium strain Gardel.";
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY652746; AAV73824.1; -; Genomic_DNA.
SQ   SEQUENCE   282 AA;  31052 MW;  0CB3C649A691BC90 CRC64;

  Query Match             76.7%;  Score 1121;  DB 2;  Length 282;
  Best Local Similarity   76.0%;  Pred. No. 7e-81;
  Matches  215;  Conservative   26;  Mismatches   38;  Indels    4;  Gaps    3;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRT--NDN-KEGFYISAKYNPSISHFRKFSA 57
                |||||||||||||||||| |||||||:|| |    |:| |||||||||||||| |||||||
Db          1 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
                ||||: | :| |||||||||:| |||  ||||     : ||| ||||||||| |||||:
Db         61 EETPVYGKDSPTKKVFGLKKEGSITKYSDFTRTDISFEGQNNFISGFSGSGIGYIMDGPRV 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
                |:|||||:|||||   |:||  :||||: ||| :|| |::|||| |:|:|| ||| | ||
Db        121 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAEAMADKKYVVLTNNGVTFSSLMFNACY 179

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
                ||||||| |:||||||||||||:|| |:||||||||||||||||||:||||||||||||
Db        180 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 239

Qy        238 GNKPEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
                |||: |:||  || | ||||:|||||||| ||||||:|:||||
Db        240 GNKYNKVPVKLPVTLIDAPQSTSASVTLDAGYFGGELGVRFTF 282


RESULT 10
Q5FGL4_EHRRG
ID   Q5FGL4_EHRRG PRELIMINARY;      PRT;   304 AA.
AC   Q5FGL4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Map1-related protein.
GN   OrderedLocusNames=ERGA_CDS_09150;
OS   Ehrlichia ruminantium (strain Gardel).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=302409;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Gardel;
RA   Frutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y.,
RA   Chantal I., Morgat A., Coissac E., Vachiery N., Demaille J., Viari A.,
RA   Martinez D.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR925677; CAI28367.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   304 AA;  33925 MW;  009338D2B65AE800 CRC64;

  Query Match             76.7%;  Score 1121;  DB 2;  Length 304;
  Best Local Similarity   76.0%;  Pred. No. 7.7e-81;
  Matches  215;  Conservative   26;  Mismatches   38;  Indels    4;  Gaps    3;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRT--NDN-KEGFYISAKYNPSISHFRKFSA 57
                |||||||||||||||||| |||||||:|| |    |:| |||||||||||||| |||||||
Db         23 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFRKFSA 82

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
                ||||: | :| |||||||||:| |||  ||||     : ||| ||||||||| |||||:
Db         83 EETPVYGKDSPTKKVFGLKKEGSITKYSDFTRTDISFEGQNNFISGFSGSGIGYIMDGPRV 142

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
                |:|||||:|||||   |:||  :||||: ||| :|| |::|||| |:|:|| ||| | ||
Db        143 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAEAMADKKYVVLTNNGVTFSSLMFNACY 201

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
                ||||||| |:||||||||||||:|| |:||||||||||||||||||:||||||||||||
Db        202 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 261

Qy        238 GNKPEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
                |||: |:||  || | ||||:|||||||| ||||||:|:||||
Db        262 GNKYNKVPVKLPVTLIDAPQSTSASVTLDAGYFGGELGVRFTF 304


RESULT 11
Q5FCD9_EHRRW
ID   Q5FCD9_EHRRW PRELIMINARY;      PRT;   304 AA.
AC   Q5FCD9;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Map1-related protein.
GN   OrderedLocusNames=ERWE_CDS_09240;
OS   Ehrlichia ruminantium (strain Welgevonden).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=254945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Welgevonden;
RA   Frutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y.,
RA   Chantal I., Morgat A., Coissac E., Vachiery N., Demaille J., Viari A.,
RA   Martinez D.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR925678; CAI27418.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   304 AA;  33925 MW;  009338D2B65AE800 CRC64;

  Query Match             76.7%;  Score 1121;  DB 2;  Length 304;
  Best Local Similarity   76.0%;  Pred. No. 7.7e-81;
  Matches  215;  Conservative   26;  Mismatches   38;  Indels    4;  Gaps    3;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRT--NDN-KEGFYISAKYNPSISHFRKFSA 57
                |||||||||||||||||| |||||||:|| |    |:| |||||||||||||| |||||||
Db         23 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFRKFSA 82

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSGIGYSMDGPRI 117
                ||||: | :| |||||||||:| |||  ||||     : ||| ||||||||| |||||:
Db         83 EETPVYGKDSPTKKVFGLKKEGSITKYSDFTRTDISFEGQNNFISGFSGSGIGYIMDGPRV 142

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
                |:|||||:|||||   |:||  :||||: ||| :|| |::|||| |:|:|| ||| | ||
Db        143 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAEAMADKKYVVLTNNGVTFSSLMFNACY 201
```

```
Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              |||||||  |:|||||||||||||:|| |:||||||||||||||||||:||||||||||||
Db        202 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 261

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||: |:||  || | ||||:|||||||| ||||||:|:||||
Db        262 GNKYNKVPVKLPVTLIDAPQSTSASVTLDAGYFGGELGVRFTF 304


RESULT 12
Q5HA10_EHRRW
ID   Q5HA10_EHRRW PRELIMINARY;      PRT;   282 AA.
AC   Q5HA10;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Putative outer membrane protein MAP1-1.
GN   Name=map1-1; OrderedLocusNames=Erum8730;
OS   Ehrlichia ruminantium (strain Welgevonden).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=254945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15637156; DOI=10.1073/pnas.0406633102;
RA   Collins N.E., Liebenberg J., de Villiers E.P., Brayton K.A., Louw E.,
RA   Pretorius A., Faber F.E., van Heerden H., Josemans A., van Kleef M.,
RA   Steyn H.C., van Strijp M.F., Zweygarth E., Jongejan F., Maillard J.C.,
RA   Berthier D., Botha M., Joubert F., Corton C.H., Thomson N.R.,
RA   Allsopp M.T., Allsopp B.A.;
RT   "The genome of the heartwater agent Ehrlichia ruminantium contains
RT   multiple tandem repeats of actively variable copy number.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:838-843(2005).
DR   EMBL; CR767821; CAH58608.1; -; Genomic_DNA.
DR   InterPro; IPR011539; RHD.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   InterPro; IPR011991; Wing_hlx_DNA_bd.
DR   Pfam; PF01617; Surface_Ag_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   282 AA;  31100 MW;  1CB2DA256B2CA990 CRC64;

  Query Match             76.6%;  Score 1120;  DB 2;  Length 282;
  Best Local Similarity   75.6%;  Pred. No. 8.4e-81;
  Matches  214;  Conservative   26;  Mismatches   39;  Indels    4;  Gaps    3;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRT--NDN-KEGFYISAKYNPSISHFRKFSA 57
              |||||||||||||||| ||||||:|| |    |:| ||||||||||||| |||||||
Db          1 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              ||||: | :| |||||||||:| |||  ||||     : ||| |||||||||| |||||:
Db         61 EETPVYGKDSPTKKVFGLKKEGSITKYSDFTRTDISFEGQNNFISGFSGSIGYIMDGPRV 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |:|||||:|||||   |:||  :||||: ||| : | |::|||| |:|:|| ||| | ||
Db        121 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAETMTDKKYVVLTNNGVTFSSLMFNACY 179

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              |||||||  |:|||||||||||||:|| |:||||||||||||||||||:||||||||||||
Db        180 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 239

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||: |:||  || | ||||:|||||||| ||||||:|:||||
Db        240 GNKYNKVPVKLPVTLIDAPQSTSASVTLDAGYFGGELGVRFTF 282


RESULT 13
Q9WW41_EHRRU
ID   Q9WW41_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q9WW41;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Major antigenic protein 1 like protein (MAP1-1).
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99216274; PubMed=10198207; DOI=10.1006/bbrc.1999.0459;
RA   Sulsona C.R., Mahan S.M., Barbet A.F.;
RT   "The map1 gene of Cowdria ruminantium is a member of a multigene
RT   family containing both conserved and variable genes.";
RL   Biochem. Biophys. Res. Commun. 257:300-305(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Welgevonden;
RA   van Heerden H., Collins N.E., Brayton K.A., Rademeyer C.,
RA   Allsopp B.A.;
RT   "Characterization of a major outer membrane protein multigene family
RT   in Ehrlichia ruminantium.";
RL   Gene 330:159-168(2004).
DR   EMBL; AF125276; AAD26347.1; -; Genomic_DNA.
DR   EMBL; AF125275; AAD26345.1; -; Genomic_DNA.
DR   EMBL; AY343331; AAR10944.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   282 AA;  31100 MW;  1CB2DA256B2CA990 CRC64;

  Query Match             76.6%;  Score 1120;  DB 2;  Length 282;
  Best Local Similarity   75.6%;  Pred. No. 8.4e-81;
  Matches  214;  Conservative   26;  Mismatches   39;  Indels    4;  Gaps    3;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRT--NDN-KEGFYISAKYNPSISHFRKFSA 57
              |||||||||||||||| ||||||:|| |    |:| ||||||||||||| |||||||
Db          1 MNYKKILVRSALISLMSFLPYQSFAEPVSSNNIGNENAKEGFYISAKYNPSIPHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              ||||: | :| |||||||||:| |||  ||||     : ||| |||||||||| |||||:
Db         61 EETPVYGKDSPTKKVFGLKKEGSITKYSDFTRTDISFEGQNNFISGFSGSIGYIMDGPRV 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |:|||||:|||||   |:||  :||||: ||| : | |::|||| |:|:|| ||| | ||
Db        121 EIEAAYQKFNPKN-PANETDTSDYYKHYGLSRAETMTDKKYVVLTNNGVTFSSLMFNACY 179

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              |||||||  |:|||||||||||||:|| |:||||||||||||||||||:||||||||||||
Db        180 DITAEGVPFIPYACAGIGADLISIFDDINLKFAYQGKIGISYPITPEISAFIGGYYHGVI 239

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              |||: |:||  || | ||||:|||||||| ||||||:|:||||
Db        240 GNKYNKVPVKLPVTLTDAPQSTSASVTLDAGYFGGELGVRFTF 282


RESULT 14
Q93DD4_EHRCH
ID   Q93DD4_EHRCH PRELIMINARY;      PRT;   275 AA.
AC   Q93DD4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Outer membrane protein p28.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=V2;
RX   MEDLINE=21893092; PubMed=11895944;
RX   DOI=10.1128/IAI.70.4.1824-1831.2002;
RA   Long S.W., Zhang X.F., Qi H., Standaert S., Walker D.H., Yu X.J.;
RT   "Antigenic variation of Ehrlichia chaffeensis resulting from
RT   differential expression of the 28-kilodalton protein gene family.";
```

```
RL   Infect. Immun. 70:1824-1831(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=V2;
RA   Yu X.-J., Zhang X.-F., Walker D.H.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF393389; AAL12919.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   275 AA;  29974 MW;  2ECCF2F988B2E9D9 CRC64;

  Query Match           45.5%;  Score 665.5;  DB 2;  Length 275;
  Best Local Similarity   49.0%;  Pred. No. 1.1e-44;
  Matches  140;  Conservative   43;  Mismatches   86;  Indels   17;  Gaps    5;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPTGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:| |            ||       ::||   ||:|:|||||||
Db         57 ----ERNTTAGVFGLKQDWDGSAIPHTQSSTPFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVN 174
              |||||| :|: |: ||  ||      | ::: ||||:  : | ::| |||:|:  :| |:|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSRQATIADNKFVFLKNEGLLDISFMLN 170

Qy        175 TCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYH 234
               |||: :||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |:||::|
Db        171 ACYDVISEGIPFSPYICAGIGTDLVSMFEVTNPKISYQGKLGLSYSISPETSVFVGGHFH 230

Qy        235 GVIGNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
                |:||:|  || | |   |       | ||||| :|| |:| || |
Db        231 KVMGNEFRDIPAIIPSGSN-LTGNHPAIVTLDVCHFGIELGGRFAF 275


RESULT 15
Q8GGU0_EHRCH
ID   Q8GGU0_EHRCH PRELIMINARY;      PRT;   276 AA.
AC   Q8GGU0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   28kDa outer membrane protein gene 19.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479838; AAO12958.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   276 AA;  30045 MW;  8E3C9719D3C67A64 CRC64;

  Query Match           44.8%;  Score 655;  DB 2;  Length 276;
  Best Local Similarity   48.8%;  Pred. No. 7.7e-44;
  Matches  140;  Conservative   43;  Mismatches   86;  Indels   18;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPTGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:| |            ||       ::||   ||:|:|||||||
Db         57 ----ERNTTAGVFGLKQDWDGSAIPHTQSSTPFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAM-EDQQYVVLKNDGITFMSLMV 173
              |||||| :|: |: ||  ||      | ::: ||||:  :  | ::| |||:|:  :| |:
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSRQATIAADNKFVFLKNEGLLDISFML 170

Qy        174 NTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYY 233
              | |||: :||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |:||::
Db        171 NACYDVISEGIPFSPYICAGIGTDLVSMFEVTNPKISYQGKLGLSYSISPETSVFVGGHF 230

Qy        234 HGVIGNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              | |:||:|  || | |   |       | ||||| :|| |:| || |
Db        231 HKVMGNEFRDIPAIIPSGSN-LTGNHPAIVTLDVCHFGIELGGRFAF 276


Search completed: January 24, 2006, 11:46:47
Job time : 212.458 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:47:54 ; Search time 2.88889 Seconds
(without alignments)
91.208 Million cell updates/sec

Title: US-10-680-349-42_COPY_61_86
Perfect score: 136
Sequence: 1 PINGTNSLTKKVFGLKKDGDITKKDD 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70609 seqs, 10134256 residues

Total number of hits satisfying chosen parameters: 70609

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	50	36.8	302	6	US-10-467-657-80	Sequence 80, Appl
2	50	36.8	302	6	US-10-467-657-4818	Sequence 4818, Ap
3	49	36.0	1992	7	US-11-013-759-3	Sequence 3, Appli
4	49	36.0	1992	7	US-11-013-759-13	Sequence 13, Appl
5	49	36.0	2047	7	US-11-013-759-4	Sequence 4, Appli
6	49	36.0	2047	7	US-11-013-759-7	Sequence 7, Appli
7	48.5	35.7	476	7	US-11-069-642-19	Sequence 19, Appl
8	48.5	35.7	585	6	US-10-510-386-20	Sequence 20, Appl
9	48.5	35.7	594	6	US-10-510-386-38	Sequence 38, Appl
10	47.5	34.9	295	6	US-10-201-525-13	Sequence 13, Appl
11	45	33.1	138	6	US-10-793-626-2878	Sequence 2878, Ap
12	45	33.1	817	6	US-10-793-626-2948	Sequence 2948, Ap
13	44	32.4	421	6	US-10-467-657-790	Sequence 790, App
14	44	32.4	433	6	US-10-793-626-1332	Sequence 1332, Ap
15	44	32.4	783	7	US-11-192-219-7	Sequence 7, Appli
16	42	30.9	416	6	US-10-641-678-73	Sequence 73, Appl
17	42	30.9	556	6	US-10-467-657-498	Sequence 498, App
18	42	30.9	1386	6	US-10-517-939-284	Sequence 284, App
19	41	30.1	129	7	US-11-125-402-24	Sequence 24, Appl
20	41	30.1	129	7	US-11-125-402-33	Sequence 33, Appl
21	41	30.1	130	7	US-11-125-402-29	Sequence 29, Appl
22	41	30.1	1431	7	US-11-128-059-2	Sequence 2, Appli
23	41	30.1	3712	7	US-11-019-711-48	Sequence 48, Appl
24	41	30.1	3712	7	US-11-019-711-51	Sequence 51, Appl
25	40.5	29.8	364	6	US-10-984-376-5	Sequence 5, Appli
26	40.5	29.8	364	6	US-10-984-376-6	Sequence 6, Appli
27	40.5	29.8	437	6	US-10-525-710-18	Sequence 18, Appl
28	40	29.4	115	6	US-10-793-626-1560	Sequence 1560, Ap
29	40	29.4	133	7	US-11-082-389-428	Sequence 428, App
30	40	29.4	295	6	US-10-467-657-8344	Sequence 8344, Ap
31	40	29.4	347	6	US-10-467-657-2962	Sequence 2962, Ap
32	40	29.4	697	7	US-11-082-389-362	Sequence 362, App
33	39	28.7	156	7	US-11-082-389-338	Sequence 338, App
34	39	28.7	229	6	US-10-656-894-1	Sequence 1, Appli
35	39	28.7	229	6	US-10-656-894-2	Sequence 2, Appli
36	39	28.7	229	6	US-10-656-894-3	Sequence 3, Appli
37	39	28.7	229	6	US-10-656-894-4	Sequence 4, Appli
38	39	28.7	309	6	US-10-793-626-356	Sequence 356, App
39	39	28.7	341	6	US-10-793-626-1278	Sequence 1278, Ap
40	39	28.7	421	7	US-11-214-199-14	Sequence 14, Appl
41	39	28.7	577	7	US-11-074-176-184	Sequence 184, App
42	39	28.7	697	7	US-11-074-176-62	Sequence 62, Appl
43	39	28.7	924	6	US-10-467-657-4290	Sequence 4290, Ap
44	39	28.7	2376	7	US-11-096-051-4	Sequence 4, Appli
45	39	28.7	2715	7	US-11-096-051-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-467-657-80
; Sequence 80, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 80
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-80

Query Match 36.8%; Score 50; DB 6; Length 302;
Best Local Similarity 57.9%; Pred. No. 1.9;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          4 GTNSLTKKVFGLKKDGDIT 22
              ||  |||  | ||||  :|
Db         27 GTIKLTKATFHLKKDTSLT 45
```

RESULT 2
US-10-467-657-4818
; Sequence 4818, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

```
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4818
;    LENGTH: 302
;    TYPE: PRT
;    ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4818

  Query Match             36.8%;  Score 50;  DB 6;  Length 302;
  Best Local Similarity   57.9%;  Pred. No. 1.9;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKDGDIT 22
              ||  |||  | ||||  :|
Db         27 GTIKLTKATFHLKKDTSLT 45


RESULT 3
US-11-013-759-3
; Sequence 3, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
;  APPLICANT: Loosmore, Sheena M.
;  APPLICANT:  Sasaki, Ken
;  APPLICANT:  Yang, Yan Ping
;  APPLICANT:  Klein, Michel H.
;  TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
;  TITLE OF INVENTION:  PROTEIN OF MORAXELLA
;  FILE REFERENCE: 1038-921MIS:jb
;  CURRENT APPLICATION NUMBER: US/11/013,759
;  CURRENT FILING DATE:  2004-12-16
;  PRIOR APPLICATION NUMBER: US/09/361,619
;  PRIOR FILING DATE: 1999-07-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;    LENGTH: 1992
;    TYPE: PRT
;    ORGANISM: Moraxella catarrhalis
US-11-013-759-3

  Query Match             36.0%;  Score 49;  DB 7;  Length 1992;
  Best Local Similarity   38.1%;  Pred. No. 26;
  Matches    8;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKDGDITKK 24
              ||  :|:   |  :|||: :|
Db        430 GTTRITRDKIGFARDGDVDEK 450


RESULT 4
US-11-013-759-13
; Sequence 13, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
;  APPLICANT: Loosmore, Sheena M.
;  APPLICANT:  Sasaki, Ken
;  APPLICANT:  Yang, Yan Ping
;  APPLICANT:  Klein, Michel H.
;  TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
;  TITLE OF INVENTION:  PROTEIN OF MORAXELLA
;  FILE REFERENCE: 1038-921MIS:jb
;  CURRENT APPLICATION NUMBER: US/11/013,759
;  CURRENT FILING DATE:  2004-12-16
;  PRIOR APPLICATION NUMBER: US/09/361,619
;  PRIOR FILING DATE: 1999-07-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;    LENGTH: 1992
;    TYPE: PRT
;    ORGANISM: Moraxella catarrhalis
US-11-013-759-13

  Query Match             36.0%;  Score 49;  DB 7;  Length 1992;
  Best Local Similarity   38.1%;  Pred. No. 26;
  Matches    8;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKDGDITKK 24
              ||  :|:   |  :|||: :|
Db        430 GTTRITRDKIGFARDGDVDEK 450


RESULT 5
US-11-013-759-4
; Sequence 4, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
;  APPLICANT: Loosmore, Sheena M.
;  APPLICANT:  Sasaki, Ken
;  APPLICANT:  Yang, Yan Ping
;  APPLICANT:  Klein, Michel H.
;  TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
;  TITLE OF INVENTION:  PROTEIN OF MORAXELLA
;  FILE REFERENCE: 1038-921MIS:jb
;  CURRENT APPLICATION NUMBER: US/11/013,759
;  CURRENT FILING DATE:  2004-12-16
;  PRIOR APPLICATION NUMBER: US/09/361,619
;  PRIOR FILING DATE: 1999-07-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;    LENGTH: 2047
;    TYPE: PRT
;    ORGANISM: Moraxella catarrhalis
US-11-013-759-4

  Query Match             36.0%;  Score 49;  DB 7;  Length 2047;
  Best Local Similarity   38.1%;  Pred. No. 27;
  Matches    8;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKDGDITKK 24
              ||  :|:   |  :|||: :|
Db        485 GTTRITRDKIGFARDGDVDEK 505


RESULT 6
US-11-013-759-7
; Sequence 7, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
;  APPLICANT: Loosmore, Sheena M.
;  APPLICANT:  Sasaki, Ken
;  APPLICANT:  Yang, Yan Ping
;  APPLICANT:  Klein, Michel H.
;  TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
;  TITLE OF INVENTION:  PROTEIN OF MORAXELLA
;  FILE REFERENCE: 1038-921MIS:jb
;  CURRENT APPLICATION NUMBER: US/11/013,759
;  CURRENT FILING DATE:  2004-12-16
;  PRIOR APPLICATION NUMBER: US/09/361,619
;  PRIOR FILING DATE: 1999-07-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;    LENGTH: 2047
;    TYPE: PRT
;    ORGANISM: Moraxella catarrhalis
US-11-013-759-7

  Query Match             36.0%;  Score 49;  DB 7;  Length 2047;
  Best Local Similarity   38.1%;  Pred. No. 27;
  Matches    8;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKDGDITKK 24
```

```
                ||  :|:   |  :|||: :|
Db        485 GTTRITRDKIGFARDGDVDEK 505
```

RESULT 7
US-11-069-642-19
; Sequence 19, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; TITLE OF INVENTION: INHIBITING PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii OT3
US-11-069-642-19

```
  Query Match             35.7%;  Score 48.5;  DB 7;  Length 476;
  Best Local Similarity   57.9%;  Pred. No. 5.4;
  Matches   11;  Conservative    4;  Mismatches    1;  Indels    3;  Gaps    1;

Qy          11 KVF---GLKKDGDITKKDD 26
               |||   |||: |:||:||:
Db          96 KVFTSEGLKEAGEITEKDE 114
```

RESULT 8
US-10-510-386-20
; Sequence 20, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-20

```
  Query Match             35.7%;  Score 48.5;  DB 6;  Length 585;
  Best Local Similarity   59.1%;  Pred. No. 7;
  Matches   13;  Conservative    2;  Mismatches    6;  Indels    1;  Gaps    1;

Qy           2 INGTNSLTKKVFG-LKKDGDIT 22
               |:||  ||||  | || :|| |
Db         432 ISGTGKLTKKGTGALKLEGDNT 453
```

RESULT 9
US-10-510-386-38
; Sequence 38, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-38

```
  Query Match             35.7%;  Score 48.5;  DB 6;  Length 594;
  Best Local Similarity   59.1%;  Pred. No. 7.1;
  Matches   13;  Conservative    2;  Mismatches    6;  Indels    1;  Gaps    1;

Qy           2 INGTNSLTKKVFG-LKKDGDIT 22
               |:||  ||||  | || :|| |
Db         437 ISGTGKLTKKGTGALKLEGDNT 458
```

RESULT 10
US-10-201-525-13
; Sequence 13, Application US/10201525
; Publication No. US20060009631A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSPERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.005
; CURRENT APPLICATION NUMBER: US/10/201,525
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-201-525-13

```
  Query Match             34.9%;  Score 47.5;  DB 6;  Length 295;
  Best Local Similarity   37.0%;  Pred. No. 4.3;
  Matches   10;  Conservative    9;  Mismatches    7;  Indels    1;  Gaps    1;

Qy           1 PINGTNSLTKKVFGLKKDGDITK-KDD 26
               | :||  ::: |: : |:||: | |:|
Db          47 PKSGTTWISEVVYMIYKEGDVEKCKED 73
```

RESULT 11
US-10-793-626-2878
; Sequence 2878, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04

```
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2878
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: amino acid sequence
US-10-793-626-2878

  Query Match            33.1%;  Score 45;  DB 6;  Length 138;
  Best Local Similarity   53.3%;  Pred. No. 4;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKD 18
              ||| ||:  ||  :|
Db         28 GTNDLTQLTFGFSRD 42


RESULT 12
US-10-793-626-2948
; Sequence 2948, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/10/793,626
;  CURRENT FILING DATE:  2004-03-04
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2948
;   LENGTH: 817
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: amino acid sequence
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (817)
;   OTHER INFORMATION: variable amino acid
US-10-793-626-2948

  Query Match            33.1%;  Score 45;  DB 6;  Length 817;
  Best Local Similarity   53.3%;  Pred. No. 35;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GTNSLTKKVFGLKKD 18
              ||| ||:  ||  :|
Db        751 GTNDLTQLTFGFSRD 765


RESULT 13
US-10-467-657-790
; Sequence 790, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 790
;   LENGTH: 421
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-657-790

  Query Match            32.4%;  Score 44;  DB 6;  Length 421;
  Best Local Similarity   36.0%;  Pred. No. 22;
  Matches    9;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKD 25
              |::|     :  |    :|||| ||
Db        301 PMDGKKEADQPDFAKDPEGDITPKD 325


RESULT 14
US-10-793-626-1332
; Sequence 1332, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/10/793,626
;  CURRENT FILING DATE:  2004-03-04
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1332
;   LENGTH: 433
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: amino acid sequence
US-10-793-626-1332

  Query Match            32.4%;  Score 44;  DB 6;  Length 433;
  Best Local Similarity   52.2%;  Pred. No. 23;
  Matches   12;  Conservative    2;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          3 NGTNSLTKKVFGLKKDGDITKKD 25
              || |   | | |  |:| |||:|
Db        140 NGVN--IKAVNGSGKNGRITKED 160


RESULT 15
US-11-192-219-7
; Sequence 7, Application US/11192219
; Publication No. US20050272656A1
;   GENERAL INFORMATION:
;        APPLICANT: Matthews, William
;        TITLE OF INVENTION: USES FOR WSX LIGANDS
;        NUMBER OF SEQUENCES: 47
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genentech, Inc.
;             STREET: 460 Point San Bruno Blvd
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: WinPatin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/11/192,219
;             FILING DATE: 27-Jul-2005
```

```
;             CLASSIFICATION: 424
;          PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/667,197
;             FILING DATE: 20-Jun-1996
;             APPLICATION NUMBER: 08/585005
;             FILING DATE: 08-Jan-96
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Lee, Wendy M.
;             REGISTRATION NUMBER: P-40,378
;             REFERENCE/DOCKET NUMBER: P0986P1
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415/225-1994
;             TELEFAX: 415/952-9881
;             TELEX: 910/371-7168
;    INFORMATION FOR SEQ ID NO: 7:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 783 amino acids
;             TYPE: Amino Acid
;             TOPOLOGY: Linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-11-192-219-7

  Query Match             32.4%;  Score 44;  DB 7;  Length 783;
  Best Local Similarity   61.5%;  Pred. No. 47;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         13 FGLKKDGDITKKD 25
              |  | |||:|||:
Db        643 FWRKMDGDVTKKE 655


Search completed: January 24, 2006, 12:04:19
Job time : 3.88889 secs
```

This Page Blank (uspto)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:46:59 ; Search time 14.3595 Seconds
(without alignments)
756.543 Million cell updates/sec

Title: US-10-680-349-42_COPY_61_86
Perfect score: 136
Sequence: 1 PINGTNSLTKKVFGLKKDGDITKKDD 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	136	100.0	280	3	US-09-811-007-42	Sequence 42, Appl
2	136	100.0	280	4	US-10-062-624-42	Sequence 42, Appl
3	136	100.0	280	4	US-10-059-964-48	Sequence 48, Appl
4	136	100.0	280	4	US-10-062-051-42	Sequence 42, Appl
5	136	100.0	280	4	US-10-062-920-42	Sequence 42, Appl
6	136	100.0	280	4	US-10-314-639-48	Sequence 48, Appl
7	136	100.0	280	4	US-10-680-349-42	Sequence 42, Appl
8	136	100.0	280	5	US-10-731-554-42	Sequence 42, Appl
9	136	100.0	280	5	US-10-901-714-48	Sequence 48, Appl
10	136	100.0	280	5	US-10-901-774-48	Sequence 48, Appl
11	97	71.3	283	3	US-09-846-808-14	Sequence 14, Appl
12	97	71.3	283	3	US-09-811-007-10	Sequence 10, Appl
13	97	71.3	283	4	US-10-062-624-10	Sequence 10, Appl
14	97	71.3	283	4	US-10-059-964-4	Sequence 4, Appli
15	97	71.3	283	4	US-10-062-051-10	Sequence 10, Appl
16	97	71.3	283	4	US-10-284-986-14	Sequence 14, Appl
17	97	71.3	283	4	US-10-062-920-10	Sequence 10, Appl
18	97	71.3	283	4	US-10-314-639-4	Sequence 4, Appli
19	97	71.3	283	4	US-10-369-293-14	Sequence 14, Appl
20	97	71.3	283	4	US-10-285-042-14	Sequence 14, Appl
21	97	71.3	283	4	US-10-680-349-10	Sequence 10, Appl
22	97	71.3	283	5	US-10-731-554-10	Sequence 10, Appl
23	97	71.3	283	5	US-10-901-714-4	Sequence 4, Appli
24	97	71.3	283	5	US-10-901-774-4	Sequence 4, Appli
25	54.5	40.1	70	4	US-10-424-599-158286	Sequence 158286,
26	53	39.0	361	4	US-10-424-599-204976	Sequence 204976,
27	52.5	38.6	339	4	US-10-369-493-4102	Sequence 4102, Ap
28	52	38.2	377	4	US-10-282-122A-57079	Sequence 57079, A
29	52	38.2	481	4	US-10-282-122A-48632	Sequence 48632, A
30	51	37.5	86	4	US-10-424-599-229045	Sequence 229045,
31	51	37.5	525	4	US-10-369-493-23031	Sequence 23031, A
32	50.5	37.1	873	4	US-10-369-493-11332	Sequence 11332, A
33	50	36.8	57	4	US-10-425-115-231193	Sequence 231193,
34	50	36.8	283	5	US-10-901-714-69	Sequence 69, Appl
35	50	36.8	283	5	US-10-901-774-69	Sequence 69, Appl
36	50	36.8	284	3	US-09-811-007-15	Sequence 15, Appl
37	50	36.8	284	4	US-10-062-624-15	Sequence 15, Appl
38	50	36.8	284	4	US-10-062-051-15	Sequence 15, Appl
39	50	36.8	284	4	US-10-062-920-15	Sequence 15, Appl
40	50	36.8	284	4	US-10-680-349-15	Sequence 15, Appl
41	50	36.8	284	5	US-10-731-554-15	Sequence 15, Appl
42	49.5	36.4	205	4	US-10-767-701-41494	Sequence 41494, A
43	49.5	36.4	346	5	US-10-771-620-6	Sequence 6, Appli
44	49.5	36.4	521	3	US-09-820-843A-93	Sequence 93, Appl
45	49.5	36.4	529	5	US-10-771-620-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-811-007-42
; Sequence 42, Application US/09811007
; Publication No. US20030185849A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-811-007-42

Query Match 100.0%; Score 136; DB 3; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 2
US-10-062-624-42
; Sequence 42, Application US/10062624
; Publication No. US20020115840A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/062,624
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42

```
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-062-624-42

  Query Match             100.0%;  Score 136;  DB 4;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86


RESULT 3
US-10-059-964-48
; Sequence 48, Application US/10059964
; Publication No. US20020120115A1
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT: Ohasi, No. US20020120115A1io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION: Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/10/059,964
;  CURRENT FILING DATE: 2002-01-28
;  EARLIER APPLICATION NUMBER: 09/314,701
;  EARLIER FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-059-964-48

  Query Match             100.0%;  Score 136;  DB 4;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86


RESULT 4
US-10-062-051-42
; Sequence 42, Application US/10062051
; Publication No. US20030073095A1
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/10/062,051
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US/09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 09/261,358
;  PRIOR FILING DATE: 1999-03-03
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-062-051-42

  Query Match             100.0%;  Score 136;  DB 4;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86


RESULT 5
US-10-062-920-42
; Sequence 42, Application US/10062920
; Publication No. US20030096250A1
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/10/062,920
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US/09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 09/261,358
;  PRIOR FILING DATE: 1999-03-03
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-062-920-42

  Query Match             100.0%;  Score 136;  DB 4;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86


RESULT 6
US-10-314-639-48
; Sequence 48, Application US/10314639
; Publication No. US20030103991A1
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT:  Ohasi, No. US20030103991A1io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION:  Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/10/314,639
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US/09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Ehrlichia canis
US-10-314-639-48

  Query Match             100.0%;  Score 136;  DB 4;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 7
US-10-680-349-42
; Sequence 42, Application US/10680349
; Publication No. US20040198951A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/680,349
; CURRENT FILING DATE: 2003-10-07
; PRIOR APPLICATION NUMBER: US/10/062,624
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-680-349-42

Query Match 100.0%; Score 136; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 8
US-10-731-554-42
; Sequence 42, Application US/10731554
; Publication No. US20040247616A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/731,554
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/811,007
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-10-731-554-42

Query Match 100.0%; Score 136; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 9
US-10-901-714-48
; Sequence 48, Application US/10901714
; Publication No. US20040265333A1
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/901,714
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
US-10-901-714-48

Query Match 100.0%; Score 136; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 10
US-10-901-774-48
; Sequence 48, Application US/10901774
; Publication No. US20040265334A1
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/901,774
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
US-10-901-774-48

Query Match 100.0%; Score 136; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 11
US-09-846-808-14
; Sequence 14, Application US/09846808
; Patent No. US20020064531A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: Yu, Xu-Jie
; TITLE OF INVENTION: Ehrlichia chaffeensis 28 kDa Outer Membrane

; TITLE OF INVENTION: Protein Multigene Family
; FILE REFERENCE: D6311
; CURRENT APPLICATION NUMBER: US/09/846,808
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,035
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 14
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: P28-14 Outer Membrane Protein of
; OTHER INFORMATION: Ehrlichia chaffeensis
US-09-846-808-14

Query Match 71.3%; Score 97; DB 3; Length 283;
Best Local Similarity 78.3%; Pred. No. 7.5e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITK 23
|||| |:||||||||||||| :
Db 64 PINGNTSITKKVFGLKKDGDIAQ 86

RESULT 12
US-09-811-007-10
; Sequence 10, Application US/09811007
; Publication No. US20030185849A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-811-007-10

Query Match 71.3%; Score 97; DB 3; Length 283;
Best Local Similarity 78.3%; Pred. No. 7.5e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITK 23
|||| |:||||||||||||| :
Db 64 PINGNTSITKKVFGLKKDGDIAQ 86

RESULT 13
US-10-062-624-10
; Sequence 10, Application US/10062624
; Publication No. US20020115840A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/062,624
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-10-062-624-10

Query Match 71.3%; Score 97; DB 4; Length 283;
Best Local Similarity 78.3%; Pred. No. 7.5e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITK 23
|||| |:||||||||||||| :
Db 64 PINGNTSITKKVFGLKKDGDIAQ 86

RESULT 14
US-10-059-964-4
; Sequence 4, Application US/10059964
; Publication No. US20020120115A1
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. US20020120115A1io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/10/059,964
; CURRENT FILING DATE: 2002-01-28
; EARLIER APPLICATION NUMBER: 09/314,701
; EARLIER FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
US-10-059-964-4

Query Match 71.3%; Score 97; DB 4; Length 283;
Best Local Similarity 78.3%; Pred. No. 7.5e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITK 23
|||| |:||||||||||||| :
Db 64 PINGNTSITKKVFGLKKDGDIAQ 86

RESULT 15
US-10-062-051-10
; Sequence 10, Application US/10062051
; Publication No. US20030073095A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/062,051
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/660,587
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B

US-10-062-051-10

Query Match 71.3%; Score 97; DB 4; Length 283;
Best Local Similarity 78.3%; Pred. No. 7.5e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86
```

Search completed: January 24, 2006, 12:03:33
Job time : 14.3595 secs

This Page Blank (uspto)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:38:28 ; Search time 5.26797 Seconds
(without alignments)
408.044 Million cell updates/sec

Title: US-10-680-349-42_COPY_61_86
Perfect score: 136
Sequence: 1 PINGTNSLTKKVFGLKKDGDITKKDD 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	136	100.0	280	2	US-09-660-587-42	Sequence 42, Appl
2	136	100.0	280	2	US-09-314-701-48	Sequence 48, Appl
3	136	100.0	280	2	US-09-811-007A-42	Sequence 42, Appl
4	136	100.0	280	2	US-10-314-639-48	Sequence 48, Appl
5	136	100.0	280	2	US-10-059-964A-48	Sequence 48, Appl
6	97	71.3	283	2	US-09-660-587-10	Sequence 10, Appl
7	97	71.3	283	2	US-09-261-358A-10	Sequence 10, Appl
8	97	71.3	283	2	US-09-201-458-6	Sequence 6, Appli
9	97	71.3	283	2	US-09-314-701-4	Sequence 4, Appli
10	97	71.3	283	2	US-09-811-007A-10	Sequence 10, Appl
11	97	71.3	283	2	US-10-314-639-4	Sequence 4, Appli
12	97	71.3	283	2	US-10-059-964A-4	Sequence 4, Appli
13	52	38.2	378	2	US-09-134-000C-5909	Sequence 5909, Ap
14	50	36.8	283	2	US-10-059-964A-69	Sequence 69, Appl
15	50	36.8	284	2	US-09-660-587-15	Sequence 15, Appl
16	50	36.8	284	2	US-09-261-358A-15	Sequence 15, Appl
17	50	36.8	284	2	US-09-201-458-11	Sequence 11, Appl
18	50	36.8	284	2	US-09-811-007A-15	Sequence 15, Appl
19	49	36.0	629	2	US-10-081-923-6	Sequence 6, Appli
20	49	36.0	1833	2	US-08-621-944A-4	Sequence 4, Appli
21	49	36.0	1833	2	US-08-945-567D-4	Sequence 4, Appli
22	49	36.0	1992	2	US-08-621-944A-3	Sequence 3, Appli
23	49	36.0	1992	2	US-08-945-567D-3	Sequence 3, Appli
24	49	36.0	2048	2	US-09-268-347-48	Sequence 48, Appl
25	48.5	35.7	416	1	US-08-464-523B-33	Sequence 33, Appl
26	48.5	35.7	476	2	US-09-800-170-19	Sequence 19, Appl
27	48	35.3	226	2	US-09-071-035-120	Sequence 120, App
28	48	35.3	226	2	US-10-206-576-120	Sequence 120, App
29	48	35.3	252	2	US-09-071-035-118	Sequence 118, App
30	48	35.3	252	2	US-10-206-576-118	Sequence 118, App
31	48	35.3	293	2	US-09-660-587-40	Sequence 40, Appl
32	48	35.3	293	2	US-09-314-701-44	Sequence 44, Appl
33	48	35.3	293	2	US-09-811-007A-40	Sequence 40, Appl
34	48	35.3	293	2	US-10-314-639-44	Sequence 44, Appl
35	48	35.3	293	2	US-10-059-964A-44	Sequence 44, Appl
36	48	35.3	300	2	US-09-314-701-50	Sequence 50, Appl
37	48	35.3	300	2	US-10-314-639-50	Sequence 50, Appl
38	48	35.3	300	2	US-10-059-964A-50	Sequence 50, Appl
39	47.5	34.9	295	2	US-09-150-133-13	Sequence 13, Appl
40	47.5	34.9	295	2	US-09-150-141-13	Sequence 13, Appl
41	47.5	34.9	295	2	US-09-374-493-13	Sequence 13, Appl
42	47.5	34.9	295	2	US-09-374-824-13	Sequence 13, Appl
43	47.5	34.9	295	2	US-09-374-492-13	Sequence 13, Appl
44	47.5	34.9	295	2	US-09-785-343-13	Sequence 13, Appl
45	47.5	34.9	295	2	US-10-411-976-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-660-587-42
; Sequence 42, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/660,587
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-660-587-42

Query Match 100.0%; Score 136; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 2
US-09-314-701-48
; Sequence 48, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT

; ORGANISM: Ehrlichia canis
US-09-314-701-48

Query Match 100.0%; Score 136; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 3
US-09-811-007A-42
; Sequence 42, Application US/09811007A
; Patent No. 6660269
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007A
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-811-007A-42

Query Match 100.0%; Score 136; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 4
US-10-314-639-48
; Sequence 48, Application US/10314639
; Patent No. 6893640
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. 6893640io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/10/314,639
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/314,701
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
US-10-314-639-48

Query Match 100.0%; Score 136; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 5
US-10-059-964A-48
; Sequence 48, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/059,964A
; CURRENT FILING DATE: 2002-01-28
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
US-10-059-964A-48

Query Match 100.0%; Score 136; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 6
US-09-660-587-10
; Sequence 10, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/660,587
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-660-587-10

Query Match 71.3%; Score 97; DB 2; Length 283;
Best Local Similarity 78.3%; Pred. No. 2.8e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITK 23
|||| |:||||||||||||| :
Db 64 PINGNTSITKKVFGLKKDGDIAQ 86

RESULT 7
US-09-261-358A-10
; Sequence 10, Application US/09261358A
; Patent No. 6403780

```
; GENERAL INFORMATION:
;  APPLICANT:  Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE:  D6152CIP
;  CURRENT APPLICATION NUMBER: US/09/261,358A
;  CURRENT FILING DATE:  1999-03-03
;  PRIOR APPLICATION NUMBER:  09/201,458
;  PRIOR FILING DATE:  1998-11-30
;  NUMBER OF SEQ ID NOS:  33
; SEQ ID NO 10
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-261-358A-10

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 2.8e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 8
US-09-201-458-6
; Sequence 6, Application US/09201458A
; Patent No. 6458942
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT: McBride, Jere W.
;  APPLICANT: Yu, Xue-Jie
;  TITLE OF INVENTION: 28-kDa Immunoreactive Protein Gene of Ehrlichia
;  TITLE OF INVENTION: canis and Uses Thereof
;  FILE REFERENCE: D6152
;  CURRENT APPLICATION NUMBER: US/09/201,458A
;  CURRENT FILING DATE: 1998-11-30
;  NUMBER OF SEQ ID NOS: 21
; SEQ ID NO 6
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-201-458-6

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 2.8e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 9
US-09-314-701-4
; Sequence 4, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT: Ohasi, No. 6544517io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION: Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/09/314,701
;  CURRENT FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-09-314-701-4

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 2.8e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 10
US-09-811-007A-10
; Sequence 10, Application US/09811007A
; Patent No. 6660269
; GENERAL INFORMATION:
;  APPLICANT: Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE: D6152CIP2
;  CURRENT APPLICATION NUMBER: US/09/811,007A
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 09/660,587
;  PRIOR FILING DATE: 2000-09-12
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 10
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B
US-09-811-007A-10

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 2.8e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 11
US-10-314-639-4
; Sequence 4, Application US/10314639
; Patent No. 6893640
; GENERAL INFORMATION:
;  APPLICANT: Rikihisa, Yasuko
;  APPLICANT:  Ohasi, No. 6893640io
;  TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
;  TITLE OF INVENTION:  Chaffeensis
;  FILE REFERENCE: 22727/04021
;  CURRENT APPLICATION NUMBER: US/10/314,639
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US/09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-10-314-639-4

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 2.8e-06;
```

```
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 12
US-10-059-964A-4
; Sequence 4, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
;  APPLICANT: RIKIHISA, YASUKO
;  APPLICANT:  OHASHI, NORIO
;  TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
;  TITLE OF INVENTION:  CHAFFEENSIS
;  FILE REFERENCE: 22727-04109
;  CURRENT APPLICATION NUMBER: US/10/059,964A
;  CURRENT FILING DATE:  2002-01-28
;  PRIOR APPLICATION NUMBER: 09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: 60/100,843
;  PRIOR FILING DATE: 1998-09-18
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Ehrlichia chaffeensis
US-10-059-964A-4

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 2.8e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 13
US-09-134-000C-5909
; Sequence 5909, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5909
;   LENGTH: 378
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-5909

  Query Match             38.2%;  Score 52;  DB 2;  Length 378;
  Best Local Similarity   39.1%;  Pred. No. 14;
  Matches    9;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 NGTNSLTKKVFGLKKDGDITKKD 25
              ||| |:|  :|  |::|::   |
Db        280 NGTKSITASLFKYKENGEVLLND 302


RESULT 14
US-10-059-964A-69
; Sequence 69, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
;  APPLICANT: RIKIHISA, YASUKO
;  APPLICANT:  OHASHI, NORIO
;  TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRLICHIA CANIS AND EHRLICHIA
;  TITLE OF INVENTION:  CHAFFEENSIS
;  FILE REFERENCE: 22727-04109
;  CURRENT APPLICATION NUMBER: US/10/059,964A
;  CURRENT FILING DATE:  2002-01-28
;  PRIOR APPLICATION NUMBER: 09/314,701
;  PRIOR FILING DATE: 1999-05-19
;  PRIOR APPLICATION NUMBER: 60/100,843
;  PRIOR FILING DATE: 1998-09-18
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 69
;   LENGTH: 283
;   TYPE: PRT
;   ORGANISM: Cowdria ruminantium
US-10-059-964A-69

  Query Match             36.8%;  Score 50;  DB 2;  Length 283;
  Best Local Similarity   73.3%;  Pred. No. 20;
  Matches   11;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 TKKVFGLKKDGDITK 23
              || ||||||| |  |
Db         63 TKAVFGLKKDWDGVK 77


RESULT 15
US-09-660-587-15
; Sequence 15, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
;  APPLICANT:  Walker, David H.
;  APPLICANT:  McBride, Jere W.
;  APPLICANT:  Yu, Xue-Jie
;  TITLE OF INVENTION:  Homologous 28-kilodalton Immunodominant Protein
;  TITLE OF INVENTION:  Genes of Ehrlichia canis and Uses Thereof
;  FILE REFERENCE:  D6152CIP2
;  CURRENT APPLICATION NUMBER: US/09/660,587
;  CURRENT FILING DATE:  2000-09-12
;  PRIOR APPLICATION NUMBER:  09/261,358
;  PRIOR FILING DATE:  1999-03-03
;  NUMBER OF SEQ ID NOS:  46
; SEQ ID NO 15
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Cowdria ruminantium
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of C. ruminantium MAP-1
US-09-660-587-15

  Query Match             36.8%;  Score 50;  DB 2;  Length 284;
  Best Local Similarity   73.3%;  Pred. No. 20;
  Matches   11;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 TKKVFGLKKDGDITK 23
              || ||||||| |  |
Db         63 TKAVFGLKKDWDGVK 77


Search completed: January 24, 2006, 11:49:00
Job time : 5.26797 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:35:23 ; Search time 19.5425 Seconds
(without alignments)
938.659 Million cell updates/sec

Title: US-10-680-349-42_COPY_61_86
Perfect score: 136
Sequence: 1 PINGTNSLTKKVFGLKKDGDITKKDD 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	136	100.0	280	2	Q84HU1_EHRCA	Q84hu1 ehrlichia c
2	136	100.0	280	2	Q9F476_EHRCA	Q9f476 ehrlichia c
3	101	74.3	283	2	Q8G8D6_EHRCH	Q8g8d6 ehrlichia c
4	101	74.3	283	2	Q8GGU4_EHRCH	Q8ggu4 ehrlichia c
5	98	72.1	283	2	Q8G8I2_EHRCH	Q8g8i2 ehrlichia c
6	97	71.3	283	2	O85358_EHRCH	O85358 ehrlichia c
7	90	66.2	282	2	Q5W5Y1_EHRRU	Q5w5y1 ehrlichia r
8	86	63.2	282	2	Q9R443_EHRRU	Q9r443 ehrlichia r
9	86	63.2	282	2	Q4L0C0_EHRRU	Q4l0c0 ehrlichia r
10	86	63.2	282	2	Q5HA10_EHRRW	Q5ha10 ehrlichia r
11	86	63.2	282	2	Q9WW41_EHRRU	Q9ww41 ehrlichia r
12	86	63.2	304	2	Q5FGL4_EHRRG	Q5fgl4 ehrlichia r
13	86	63.2	304	2	Q5FCD9_EHRRW	Q5fcd9 ehrlichia r
14	64	47.1	238	2	Q8A3S3_BACTN	Q8a3s3 bacteroides
15	57	41.9	266	2	Q7R7D4_PLAYO	Q7r7d4 plasmodium
16	57	41.9	266	2	Q7RER7_PLAYO	Q7rer7 plasmodium
17	57	41.9	268	2	Q7RA08_PLAYO	Q7ra08 plasmodium
18	55.5	40.8	481	2	Q5NHN7_FRATT	Q5nhn7 francisella
19	55	40.4	92	2	Q98J82_RHILO	Q98j82 rhizobium l
20	54	39.7	1333	2	Q54KL3_DICDI	Q54kl3 dictyosteli
21	53.5	39.3	2515	2	O77365_PLAF7	O77365 plasmodium
22	53	39.0	266	2	Q7REB5_PLAYO	Q7reb5 plasmodium
23	53	39.0	426	1	GLYA_THEVO	Q97cq5 thermoplasm
24	53	39.0	800	2	Q6BNK2_DEBHA	Q6bnk2 debaryomyce
25	53	39.0	1730	2	Q7RDQ0_PLAYO	Q7rdq0 plasmodium
26	52.5	38.6	324	2	Q7S7U0_NEUCR	Q7s7u0 neurospora
27	52	38.2	377	2	Q8KU85_ENTFA	Q8ku85 enterococcu
28	52	38.2	481	2	Q5L9U4_BACFN	Q5l9u4 bacteroides
29	52	38.2	481	2	Q64Q48_BACFR	Q64q48 bacteroides
30	52	38.2	912	2	Q5L9D8_BACFN	Q5l9d8 bacteroides
31	52	38.2	912	2	Q64PM2_BACFR	Q64pm2 bacteroides
32	52	38.2	1218	2	Q4P674_USTMA	Q4p674 ustilago ma
33	51.5	37.9	279	2	Q6NWI8_BRARE	Q6nwi8 brachydanio
34	51.5	37.9	551	2	Q8I637_PLAF7	Q8i637 plasmodium
35	51.5	37.9	562	2	Q6C527_YARLI	Q6c527 yarrowia li
36	51	37.5	321	2	Q6CFV8_YARLI	Q6cfv8 yarrowia li
37	51	37.5	345	2	Q93FL9_9ENTR	Q93fl9 citrobacter
38	51	37.5	530	2	Q5KW58_GEOKA	Q5kw58 geobacillus
39	51	37.5	559	2	Q9FRS4_ARATH	Q9frs4 arabidopsis
40	51	37.5	948	2	Q84N32_SORBI	Q84n32 sorghum bic
41	51	37.5	1373	2	Q7RQ53_PLAYO	Q7rq53 plasmodium
42	51	37.5	1793	2	Q8EWW0_MYCPE	Q8eww0 mycoplasma
43	50.5	37.1	268	2	Q7YTF3_CRYPV	Q7ytf3 cryptospori
44	50.5	37.1	281	2	Q4QWR1_CRYPV	Q4qwr1 cryptospori
45	50.5	37.1	282	2	Q4QWR0_CRYPV	Q4qwr0 cryptospori

ALIGNMENTS

RESULT 1
Q84HU1_EHRCA
ID Q84HU1_EHRCA PRELIMINARY; PRT; 280 AA.
AC Q84HU1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major outer membrane protein.
GN Name=p30-10;
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=944;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Hawaii;
RX MEDLINE=22461952; PubMed=12574308; DOI=10.1128/JCM.41.2.886-888.2003;
RA Felek S., Greene R., Rikihisa Y.;
RT "Transcriptional Analysis of p30 Major Outer Membrane Protein Genes of
RT Ehrlichia canis in Naturally Infected Ticks and Sequence Analysis of
RT p30-10 of E. canis from Diverse Geographic Regions.";
RL J. Clin. Microbiol. 41:886-888(2003).
DR EMBL; AF528513; AAO41110.1; -; Genomic_DNA.
DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 280 AA; 30990 MW; D1BB28B14F5BDCA2 CRC64;

Query Match 100.0%; Score 136; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 2
Q9F476_EHRCA
ID Q9F476_EHRCA PRELIMINARY; PRT; 280 AA.
AC Q9F476;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE P28-2 (Major outer membrane protein P30-10).
GN Name=p28-2; Synonyms=p30-10;
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=944;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Jake;
RX MEDLINE=99242757; PubMed=10225842;
RA McBride J.W., Yu, Xj, Walker D.H.;

```
RT   "Molecular cloning of the gene for a conserved major immunoreactive
RT   28-kilodalton protein of Ehrlichia canis: a potential serodiagnostic
RT   antigen.";
RL   Clin. Diagn. Lab. Immunol. 6:392-399(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Jake;
RX   MEDLINE=20432107; PubMed=10974556; DOI=10.1016/S0378-1119(00)00256-0;
RA   McBride J.W., Yu X.J., Walker D.H.;
RT   "A conserved, transcriptionally active p28 multigene locus of
RT   Ehrlichia canis.";
RL   Gene 254:245-252(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Oklahoma;
RX   MEDLINE=98371112; PubMed=9705412;
RA   Ohashi N., Unver A., Zhi N., Rikihisa Y.;
RT   "Cloning and characterization of multigenes encoding the
RT   immunodominant 30-kilodalton major outer membrane proteins of
RT   Ehrlichia canis and application of the recombinant protein for
RT   serodiagnosis.";
RL   J. Clin. Microbiol. 36:2671-2680(1998).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Oklahoma;
RX   MEDLINE=21153566; PubMed=11254561;
RX   DOI=10.1128/IAI.69.4.2083-2091.2001;
RA   Ohashi N., Rikihisa Y., Unver A.;
RT   "Analysis of transcriptionally active gene clusters of major outer
RT   membrane protein multigene family in Ehrlichia canis and E.
RT   chaffeensis.";
RL   Infect. Immun. 69:2083-2091(2001).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Venezuela, Arizona, California, and New Mexico;
RX   MEDLINE=22461952; PubMed=12574308; DOI=10.1128/JCM.41.2.886-888.2003;
RA   Felek S., Greene R., Rikihisa Y.;
RT   "Transcriptional Analysis of p30 Major Outer Membrane Protein Genes of
RT   Ehrlichia canis in Naturally Infected Ticks and Sequence Analysis of
RT   p30-10 of E. canis from Diverse Geographic Regions.";
RL   J. Clin. Microbiol. 41:886-888(2003).
DR   EMBL; AF082744; AAG14357.1; -; Genomic_DNA.
DR   EMBL; AF078553; AAK28696.1; -; Genomic_DNA.
DR   EMBL; AF324792; AAK31312.1; -; Genomic_DNA.
DR   EMBL; AF528512; AAO41109.1; -; Genomic_DNA.
DR   EMBL; AF528515; AAO41112.1; -; Genomic_DNA.
DR   EMBL; AF528514; AAO41111.1; -; Genomic_DNA.
DR   EMBL; AF528511; AAO41108.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   280 AA;  30962 MW;  D1BB28B5AF0E6CA3 CRC64;

  Query Match            100.0%;  Score 136;  DB 2;  Length 280;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86


RESULT 3
Q8G8D6_EHRCH
ID   Q8G8D6_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   Q8G8D6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   28kDa outer membrane protein gene 14.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479835; AAO12939.1; -; Genomic_DNA.
DR   EMBL; AF479836; AAO12944.1; -; Genomic_DNA.
DR   EMBL; AF479838; AAO12954.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31090 MW;  D5825DD4DC51C425 CRC64;

  Query Match            74.3%;  Score 101;  DB 2;  Length 283;
  Best Local Similarity  82.6%;  Pred. No. 2.5e-06;
  Matches   19;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||| :|||||||||||||| :
Db         64 PINGTTALTKKVFGLKKDGDIAQ 86


RESULT 4
Q8GGU4_EHRCH
ID   Q8GGU4_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   Q8GGU4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   28kDa outer membrane protein gene 14.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479837; AAO12949.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31042 MW;  A2960DD4DC45942A CRC64;

  Query Match            74.3%;  Score 101;  DB 2;  Length 283;
  Best Local Similarity  82.6%;  Pred. No. 2.5e-06;
  Matches   19;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||| :|||||||||||||| :
Db         64 PINGTTALTKKVFGLKKDGDIAQ 86


RESULT 5
Q8G8I2_EHRCH
ID   Q8G8I2_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   Q8G8I2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   28kDa outer membrane protein gene 14.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
```

```
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
DR   EMBL; AF479839; AAO12959.1; -; Genomic_DNA.
DR   EMBL; AF479840; AAO12965.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31069 MW;  AEA8B2833E02631F CRC64;

  Query Match             72.1%;  Score 98;  DB 2;  Length 283;
  Best Local Similarity   69.2%;  Pred. No. 7e-06;
  Matches   18;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||  ::|||||||||||||  :  |
Db         64 PINGNTAITKKVFGLKKDGDIAQHGD 89


RESULT 6
O85358_EHRCH
ID   O85358_EHRCH PRELIMINARY;      PRT;   283 AA.
AC   O85358; O52103;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   28kDa outer membrane protein gene 14 (Major outer membrane protein
DE   OMP-1B).
GN   Name=omp-1B;
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=98321180; PubMed=9647746; DOI=10.1006/bbrc.1998.8844;
RA   Reddy G.R., Sulsona C.R., Barbet A.F., Mahan S.M., Burridge M.J.,
RA   Alleman A.R.;
RT   "Molecular characterization of a 28 kDa surface antigen gene family of
RT   the tribe Ehrlichiae.";
RL   Biochem. Biophys. Res. Commun. 247:636-643(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22384137; PubMed=12496165; DOI=10.1128/IAI.71.1.187-195.2003;
RA   Cheng C., Paddock C.D., Ganta R.R.;
RT   "Molecular heterogeneity of Ehrlichia chaffeensis isolates determined
RT   by sequence analysis of the 28-kilodalton outer membrane protein genes
RT   and other regions of the genome.";
RL   Infect. Immun. 71:187-195(2003).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Cheng C., Paddock C.D., Ganta R.R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Arkansas;
RX   MEDLINE=98084465; PubMed=9423849;
RA   Ohashi N., Zhi N., Zhang Y., Rikihisa Y.;
RT   "Immunodominant major outer membrane proteins of Ehrlichia chaffeensis
RT   are encoded by a polymorphic multigene family.";
RL   Infect. Immun. 66:132-139(1998).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Arkansas;
RX   MEDLINE=21153566; PubMed=11254561;
RX   DOI=10.1128/IAI.69.4.2083-2091.2001;
RA   Ohashi N., Rikihisa Y., Unver A.;
RT   "Analysis of transcriptionally active gene clusters of major outer
RT   membrane protein multigene family in Ehrlichia canis and E.
RT   chaffeensis.";
RL   Infect. Immun. 69:2083-2091(2001).
DR   EMBL; AF479833; AAO12929.1; -; Genomic_DNA.
DR   EMBL; U72291; AAC02936.1; -; Genomic_DNA.
DR   EMBL; AF479834; AAO12933.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   283 AA;  31017 MW;  DCBCF652B771C95D CRC64;

  Query Match             71.3%;  Score 97;  DB 2;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 9.8e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||  :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 7
Q5W5Y1_EHRRU
ID   Q5W5Y1_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q5W5Y1;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   MAP1-1.
GN   Name=map1-1;
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Senegal;
RX   MEDLINE=22035368; PubMed=12039046; DOI=10.1016/S0378-1119(02)00408-0;
RA   Bekker C.P., Bell-Sakyi L., Paxton E.A., Martinez D., Bensaid A.,
RA   Jongejan F.;
RT   "Transcriptional analysis of the major antigenic protein 1 multigene
RT   family of Cowdria ruminantium.";
RL   Gene 285:193-201(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Senegal;
RX   PubMed=15995193; DOI=10.1128/JB.187.14.4782-4791.2005;
RA   Bekker C.P., Postigo M., Taoufik A., Bell-Sakyi L., Ferraz C.,
RA   Martinez D., Jongejan F.;
RT   "Transcription Analysis of the Major Antigenic Protein 1 Multigene
RT   Family of Three In Vitro-Cultured Ehrlichia ruminantium Isolates.";
RL   J. Bacteriol. 187:4782-4791(2005).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Senegal;
RA   Bekker C.P.J., Taoufik A., Jongejan F.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF319940; AAV54088.2; -; Genomic_DNA.
SQ   SEQUENCE   282 AA;  31097 MW;  3D619510E7393EBB CRC64;

  Query Match             66.2%;  Score 90;  DB 2;  Length 282;
  Best Local Similarity   69.2%;  Pred. No. 0.00011;
  Matches   18;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| |||||||||||  |||  |
Db         64 PVYGKDSPTKKVFGLKKDGSITKYSD 89


RESULT 8
Q9R443_EHRRU
ID   Q9R443_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q9R443;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Major antigenic protein 1 like protein.
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
```

```
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99216274; PubMed=10198207; DOI=10.1006/bbrc.1999.0459;
RA   Sulsona C.R., Mahan S.M., Barbet A.F.;
RT   "The map1 gene of Cowdria ruminantium is a member of a multigene
RT   family containing both conserved and variable genes.";
RL   Biochem. Biophys. Res. Commun. 257:300-305(1999).
DR   EMBL; AF125279; AAD26353.1; -; Genomic_DNA.
DR   EMBL; AF125274; AAD26343.1; -; Genomic_DNA.
DR   EMBL; AF125277; AAD26349.1; -; Genomic_DNA.
DR   EMBL; AF125278; AAD26351.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   282 AA;  31040 MW;  0CB3C655775CBC90 CRC64;

  Query Match             63.2%;  Score 86;  DB 2;  Length 282;
  Best Local Similarity   65.4%;  Pred. No. 0.00041;
  Matches   17;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| ||||||||:| |||  |
Db         64 PVYGKDSPTKKVFGLKKEGSITKYSD 89


RESULT 9
Q4L0C0_EHRRU
ID   Q4L0C0_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q4L0C0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Map1-1.
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Gardel;
RA   Martinez D., Ferraz C., Demaille J., Bensaid A., Frutos R.;
RT   "Analysis of the major antigenic protein 1 multigene family of
RT   Ehrlichia ruminantium strain Gardel.";
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY652746; AAV73824.1; -; Genomic_DNA.
SQ   SEQUENCE   282 AA;  31052 MW;  0CB3C649A691BC90 CRC64;

  Query Match             63.2%;  Score 86;  DB 2;  Length 282;
  Best Local Similarity   65.4%;  Pred. No. 0.00041;
  Matches   17;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| ||||||||:| |||  |
Db         64 PVYGKDSPTKKVFGLKKEGSITKYSD 89


RESULT 10
Q5HA10_EHRRW
ID   Q5HA10_EHRRW PRELIMINARY;      PRT;   282 AA.
AC   Q5HA10;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Putative outer membrane protein MAP1-1.
GN   Name=map1-1; OrderedLocusNames=Erum8730;
OS   Ehrlichia ruminantium (strain Welgevonden).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=254945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15637156; DOI=10.1073/pnas.0406633102;
RA   Collins N.E., Liebenberg J., de Villiers E.P., Brayton K.A., Louw E.,
RA   Pretorius A., Faber F.E., van Heerden H., Josemans A., van Kleef M.,
RA   Steyn H.C., van Strijp M.F., Zweygarth E., Jongejan F., Maillard J.C.,
RA   Berthier D., Botha M., Joubert F., Corton C.H., Thomson N.R.,
RA   Allsopp M.T., Allsopp B.A.;
RT   "The genome of the heartwater agent Ehrlichia ruminantium contains
RT   multiple tandem repeats of actively variable copy number.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:838-843(2005).
DR   EMBL; CR767821; CAH58608.1; -; Genomic_DNA.
DR   InterPro; IPR011539; RHD.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   InterPro; IPR011991; Wing_hlx_DNA_bd.
DR   Pfam; PF01617; Surface_Ag_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   282 AA;  31100 MW;  1CB2DA256B2CA990 CRC64;

  Query Match             63.2%;  Score 86;  DB 2;  Length 282;
  Best Local Similarity   65.4%;  Pred. No. 0.00041;
  Matches   17;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| ||||||||:| |||  |
Db         64 PVYGKDSPTKKVFGLKKEGSITKYSD 89


RESULT 11
Q9WW41_EHRRU
ID   Q9WW41_EHRRU PRELIMINARY;      PRT;   282 AA.
AC   Q9WW41;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Major antigenic protein 1 like protein (MAP1-1).
OS   Ehrlichia ruminantium (Cowdria ruminantium).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=779;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99216274; PubMed=10198207; DOI=10.1006/bbrc.1999.0459;
RA   Sulsona C.R., Mahan S.M., Barbet A.F.;
RT   "The map1 gene of Cowdria ruminantium is a member of a multigene
RT   family containing both conserved and variable genes.";
RL   Biochem. Biophys. Res. Commun. 257:300-305(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Welgevonden;
RA   van Heerden H., Collins N.E., Brayton K.A., Rademeyer C.,
RA   Allsopp B.A.;
RT   "Characterization of a major outer membrane protein multigene family
RT   in Ehrlichia ruminantium.";
RL   Gene 330:159-168(2004).
DR   EMBL; AF125276; AAD26347.1; -; Genomic_DNA.
DR   EMBL; AF125275; AAD26345.1; -; Genomic_DNA.
DR   EMBL; AY343331; AAR10944.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   282 AA;  31100 MW;  1CB2DA256B2CA990 CRC64;

  Query Match             63.2%;  Score 86;  DB 2;  Length 282;
  Best Local Similarity   65.4%;  Pred. No. 0.00041;
  Matches   17;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| ||||||||:| |||  |
Db         64 PVYGKDSPTKKVFGLKKEGSITKYSD 89


RESULT 12
Q5FGL4_EHRRG
ID   Q5FGL4_EHRRG PRELIMINARY;      PRT;   304 AA.
AC   Q5FGL4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
```

```
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Map1-related protein.
GN   OrderedLocusNames=ERGA_CDS_09150;
OS   Ehrlichia ruminantium (strain Gardel).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=302409;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Gardel;
RA   Frutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y.,
RA   Chantal I., Morgat A., Coissac E., Vachiery N., Demaille J., Viari A.,
RA   Martinez D.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR925677; CAI28367.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   304 AA;  33925 MW;  009338D2B65AE800 CRC64;

  Query Match             63.2%;  Score 86;  DB 2;  Length 304;
  Best Local Similarity   65.4%;  Pred. No. 0.00044;
  Matches   17;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| |||||||||:| |||  |
Db         86 PVYGKDSPTKKVFGLKKEGSITKYSD 111


RESULT 13
Q5FCD9_EHRRW
ID   Q5FCD9_EHRRW PRELIMINARY;      PRT;   304 AA.
AC   Q5FCD9;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Map1-related protein.
GN   OrderedLocusNames=ERWE_CDS_09240;
OS   Ehrlichia ruminantium (strain Welgevonden).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=254945;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Welgevonden;
RA   Frutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y.,
RA   Chantal I., Morgat A., Coissac E., Vachiery N., Demaille J., Viari A.,
RA   Martinez D.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR925678; CAI27418.1; -; Genomic_DNA.
DR   InterPro; IPR002566; Surface_Ag_msp4.
DR   Pfam; PF01617; Surface_Ag_2; 1.
SQ   SEQUENCE   304 AA;  33925 MW;  009338D2B65AE800 CRC64;

  Query Match             63.2%;  Score 86;  DB 2;  Length 304;
  Best Local Similarity   65.4%;  Pred. No. 0.00044;
  Matches   17;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |: | :| |||||||||:| |||  |
Db         86 PVYGKDSPTKKVFGLKKEGSITKYSD 111


RESULT 14
Q8A3S3_BACTN
ID   Q8A3S3_BACTN PRELIMINARY;      PRT;   238 AA.
AC   Q8A3S3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase.
GN   OrderedLocusNames=BT2881;
OS   Bacteroides thetaiotaomicron.
OC   Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC   Bacteroidaceae; Bacteroides.
OX   NCBI_TaxID=818;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=VPI-5482 / ATCC 29148;
RX   MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA   Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA   Chiang H.C., Hooper L.V., Gordon J.I.;
RT   "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL   Science 299:2074-2076(2003).
DR   EMBL; AE016937; AAO77987.1; -; Genomic_DNA.
DR   GO; GO:0050518; F:2-C-methyl-D-erythritol 4-phosphate cytidyl. . .; IEA.
DR   GO; GO:0008699; F:4-diphosphocytidyl-2C-methyl-D-erythritol s. . .; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0008299; P:isoprenoid biosynthesis; IEA.
DR   InterPro; IPR008233; DPCME_synth.
DR   InterPro; IPR001228; ISPD_synthase.
DR   Pfam; PF01128; IspD; 1.
DR   PIRSF; PIRSF006765; DPCME_synth; 1.
DR   PROSITE; PS01295; ISPD; UNKNOWN_1.
KW   Complete proteome; Nucleotidyltransferase; Transferase.
SQ   SEQUENCE   238 AA;  26323 MW;  8C5685CFBE8DE768 CRC64;

  Query Match             47.1%;  Score 64;  DB 2;  Length 238;
  Best Local Similarity   52.0%;  Pred. No. 0.6;
  Matches   13;  Conservative    5;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 INGTNSLTKKVFGLKKDGDITKKDD 26
              |||  |::  ::||||:|   ||||
Db         80 INGQESISNGIYGLKKNG--VKKDD 102


RESULT 15
Q7R7D4_PLAYO
ID   Q7R7D4_PLAYO PRELIMINARY;      PRT;   266 AA.
AC   Q7R7D4;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative yir3 protein.
GN   Name=PY07653;
OS   Plasmodium yoelii yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=73239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=17XNL;
RX   MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA   Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA   Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA   Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA   Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA   Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA   Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA   Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA   van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA   Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA   Carucci D.J.;
RT   "Genome sequence and comparative analysis of the model rodent malaria
RT   parasite Plasmodium yoelii yoelii.";
RL   Nature 419:512-519(2002).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABL01002851; EAA20157.1; -; Genomic_DNA.
DR   InterPro; IPR006477; Yir_bir_cir.
DR   Pfam; PF06022; Cir_Bir_Yir; 1.
DR   TIGRFAMs; TIGR01590; yir-bir-cir_Pla; 1.
SQ   SEQUENCE   266 AA;  31613 MW;  AC785D5A383F8C5E CRC64;

  Query Match             41.9%;  Score 57;  DB 2;  Length 266;
```

```
  Best Local Similarity   50.0%;  Pred. No. 7.2;
  Matches   12;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 INGTNSLTKKVFGLKKDGDITKKD 25
              : | |   ||   |||| || | |
Db        165 LEGDNEFVKKYDQLKKDSDINKDD 188


Search completed: January 24, 2006, 11:46:49
Job time : 21.5425 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:36:23 ; Search time 4.07843 Seconds
(without alignments)
613.382 Million cell updates/sec

Title: US-10-680-349-42_COPY_61_86
Perfect score: 136
Sequence: 1 PINGTNSLTKKVFGLKKDGDITKKDD 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	53.5	39.3	2523	2	T18477	hypothetical prote
2	51	37.5	559	2	T00714	hypothetical prote
3	50	36.8	284	2	I40882	major antigenic pr
4	49.5	36.4	521	2	B71717	hypothetical prote
5	49.5	36.4	1650	2	T18444	hypothetical prote
6	49	36.0	345	2	D91201	type III secretion
7	49	36.0	345	2	H86047	escU [imported) -
8	49	36.0	363	2	F97092	probable membrane
9	49	36.0	1245	2	S51255	probable membrane
10	48.5	35.7	608	2	T02299	hypothetical prote
11	48.5	35.7	1092	1	JN0635	neural cell adhesi
12	48.5	35.7	1127	1	E71156	endopeptidase La h
13	48	35.3	155	2	T29928	hypothetical prote
14	48	35.3	274	2	H83707	bacitracin resista
15	48	35.3	427	2	F75169	serine hydroxymeth
16	48	35.3	635	2	B72215	hypothetical prote
17	48	35.3	900	2	G87431	pyruvate phosphate
18	48	35.3	908	2	AE2675	pyruvate,orthophos
19	48	35.3	930	2	AF3431	pyruvate, phosphat
20	48	35.3	933	2	C97457	pyruvate, phosphat
21	47.5	34.9	295	2	I53296	testis-specific es
22	47.5	34.9	564	2	A34325	juvenile-hormone e
23	47.5	34.9	881	2	F72397	pyruvate, phosphat
24	47	34.6	427	2	F71045	probable serine hy
25	46.5	34.2	833	2	H72205	maltose ABC transp
26	46.5	34.2	879	2	AC1308	pyruvate phosphate
27	46.5	34.2	879	2	AC1680	pyruvate phosphate
28	46.5	34.2	947	1	KIZMPO	pyruvate, phosphat
29	46.5	34.2	947	2	T02979	pyruvate, phosphat
30	46.5	34.2	949	1	S55478	pyruvate, phosphat
31	46.5	34.2	953	2	S56650	pyruvate, phosphat
32	46.5	34.2	953	2	S12894	pyruvate, phosphat
33	46.5	34.2	955	2	S56649	pyruvate, phosphat
34	46.5	34.2	956	1	S53297	pyruvate, phosphat
35	46.5	34.2	960	2	A71420	pyruvate, phosphat
36	46	33.8	225	2	T21094	hypothetical prote
37	46	33.8	440	2	T01115	probable protein d
38	46	33.8	517	2	A71707	lipopolysaccharide
39	46	33.8	631	2	S70908	transferrin-bindin
40	46	33.8	840	1	KIQAPO	pyruvate, phosphat
41	46	33.8	1123	2	T28139	PK4 protein kinase
42	46	33.8	1505	2	T31418	synaptonemal compl
43	46	33.8	2469	2	H36812	hypothetical prote
44	45.5	33.5	70	2	B64497	hypothetical prote
45	45.5	33.5	72	2	C95929	hypothetical expor

ALIGNMENTS

RESULT 1
T18477
hypothetical protein C0485w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18477
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A;Reference number: Z18937
A;Accession: T18477
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2523 <LAW>
A;Cross-references: UNIPROT:O77365; UNIPARC:UPI000017CC31; EMBL:AL008970; NID:e1407852;
C;Genetics:
A;Map position: 3
A;Introns: 148/3
A;Note: C0485w

Query Match 39.3%; Score 53.5; DB 2; Length 2523;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 15; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

```
Qy          2 INGTNSLTKKVFGLKKD-----GDITKKDD 26
              ||  | : ||    |||     ||| ||||
Db       1822 INKKNDINKKDDINKKDDINKKGDINKKDD 1851
```

RESULT 2
T00714
hypothetical protein F22O13.9 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00714
R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Conv
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, April 1998
A;Description: Genomic sequence for Arabidopsis thaliana BAC F22O13.
A;Reference number: Z14200
A;Accession: T00714
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-559 <SHI>
A;Cross-references: UNIPROT:Q9FRS4; UNIPARC:UPI00000A4F36; EMBL:AC003981; NID:g3063438;
C;Genetics:
A;Gene: ATSP:F22O13.9
A;Map position: 1

Query Match 37.5%; Score 51; DB 2; Length 559;
Best Local Similarity 44.4%; Pred. No. 21;
Matches 12; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

```
Qy          4 GTNSLTK----KVFGLKKDGDITKKDD 26
              |||    :    | |||  || ||: |:
Db         78 GTNLRARVKPMKQFGLSSDGPITENDE 104


RESULT 3
I40882
major antigenic protein - heartwater rickettsia
C;Species: Cowdria ruminantium (heartwater rickettsia)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40882; S42827
R;van Vliet, A.H.; Jongejan, F.; van Kleef, M.; van der Zeijst, B.A.
Infect. Immun. 62, 1451-1456, 1994
A;Title: Molecular cloning, sequence analysis, and expression of the gene encoding the i
A;Reference number: I40882; MUID:94178956; PMID:8132352
A;Accession: I40882
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-284 <RES>
A;Cross-references: UNIPROT:Q46327; UNIPARC:UPI00000B9A44; EMBL:X74250; NID:g454266; PID
C;Genetics:
A;Gene: map1

  Query Match             36.8%;  Score 50;  DB 2;  Length 284;
  Best Local Similarity   73.3%;  Pred. No. 14;
  Matches   11;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 TKKVFGLKKDGDITK 23
              || ||||||| |  |
Db         63 TKAVFGLKKDWDGVK 77


RESULT 4
B71717
hypothetical protein RP083 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: B71717
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: B71717
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-521 <AND>
A;Cross-references: UNIPROT:Q9ZE64; UNIPARC:UPI00000C114A; GB:AJ235270; GB:AJ235269; NID
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: RP083

  Query Match             36.4%;  Score 49.5;  DB 2;  Length 521;
  Best Local Similarity   47.8%;  Pred. No. 32;
  Matches   11;  Conservative    5;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          3 NGTNSLTKKVFGLKKDGDITKKD 25
              |  | : :|:| || || ||::|
Db         83 NLNNKILEKIF-LKDDGTITEQD 104


RESULT 5
T18444
hypothetical protein C0385c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18444
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z18935
A;Accession: T18444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1650 <LAW>
A;Cross-references: UNIPROT:O77328; UNIPARC:UPI00000748ED; EMBL:Z98547; NID:e1325376; PII
C;Genetics:
A;Map position: 3
A;Introns: 1597/3; 1625/3
A;Note: C0385c

  Query Match             36.4%;  Score 49.5;  DB 2;  Length 1650;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches   12;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          6 NSLTKKVFGLKKDGDITKKDD 26
              ::| |||| :|||  ||  ||
Db        932 STLEKKVF-VKKDNVITNNDD 951


RESULT 6
D91201
type III secretion system EscU protein [imported] - Escherichia coli (strain O157:H7, sub
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D91201
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genon
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D91201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-345 <HAY>
A;Cross-references: UNIPROT:O85640; UNIPARC:UPI00000D00D0; GB:BA000007; PIDN:BAB38003.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4580
C;Superfamily: flagellar biosynthetic protein flhB; flhB carboxyl-terminal homology

  Query Match             36.0%;  Score 49;  DB 2;  Length 345;
  Best Local Similarity   52.9%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 KKVFGLKKDGDITKKDD 26
              ||:  ||| ||:|| ::
Db         11 KKLRDLKKKGDVTKSEE 27


RESULT 7
H86047
escU [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: H86047
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H86047
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-345 <STO>
A;Cross-references: UNIPROT:O85640; UNIPARC:UPI00000D00D0; GB:AE005174; NID:g12518474; PI
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: escU
C;Superfamily: flagellar biosynthetic protein flhB; flhB carboxyl-terminal homology

  Query Match             36.0%;  Score 49;  DB 2;  Length 345;
  Best Local Similarity   52.9%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 KKVFGLKKDGDITKKDD 26
              ||:  ||| ||:|| ::
Db         11 KKLRDLKKKGDVTKSEE 27
```

```
RESULT 8
F97092
probable membrane protein [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97092
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: F97092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-363 <KUR>
A;Cross-references: UNIPROT:Q97IS8; UNIPARC:UPI00000CA244; GB:AE001437; PIDN:AAK79529.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1562

  Query Match             36.0%;  Score 49;  DB 2;  Length 363;
  Best Local Similarity   50.0%;  Pred. No. 26;
  Matches   11;  Conservative    3;  Mismatches    4;  Indels    4;  Gaps    1;

Qy          9 TKKVF----GLKKDGDITKKDD 26
              | ||:    ||:||||:  | |
Db        152 TGKVYIPLDGLEKDGDVKNKTD 173


RESULT 9
S51255
probable membrane protein YDR104c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YD8557.13c
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S51255
R;Murphy, L.; Harris, D.
submitted to the EMBL Data Library, January 1995
A;Reference number: S51243
A;Accession: S51255
A;Molecule type: DNA
A;Residues: 1-1245 <MUR>
A;Cross-references: UNIPROT:Q03868; UNIPARC:UPI000006B2F3; EMBL:Z47746; NID:g633627; PID
C;Genetics:
A;Gene: SGD:SPO71; MIPS:YDR104c
A;Cross-references: SGD:S0002511
A;Map position: 4R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YDR104c
C;Keywords: transmembrane protein
F;1045-1061/Domain: transmembrane #status predicted <TMM>

  Query Match             36.0%;  Score 49;  DB 2;  Length 1245;
  Best Local Similarity   52.4%;  Pred. No. 95;
  Matches   11;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 TNSLTKKVFGLKKDGDITKKD 25
              ||||| |    |: | | ||:
Db        440 TNSLTNKATVKKRTGQILKKE 460


RESULT 10
T02299
hypothetical protein F23858_1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02299; T46353
R;Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Adamson, A.W.; Burkhart-Schultz, K.; G
.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Attix, C.; Andreise, T.; Trankhei
submitted to the EMBL Data Library, March 1998
A;Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Arell
A;Description: Sequence analysis of a human P1 clone containing the XRCC9 DNA repair gen
A;Reference number: Z14637
A;Accession: T02299
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-608 <LAM>
A;Cross-references: UNIPROT:O60378; UNIPARC:UPI0000073023; EMBL:AC004475; NID:g2988396; I
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23037
A;Accession: T46353
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 217-608 <AAA>
A;Cross-references: UNIPARC:UPI000007056F; EMBL:AL137286
A;Experimental source: adult testis; clone DKFZp434E2216
C;Genetics:
A;Map position: 19
A;Introns: 58/1; 84/1; 143/3; 185/1; 218/3; 260/1; 378/3; 414/2; 473/3; 509/2; 558/1; 601
A;Note: DKFZp434E2216.1

  Query Match             35.7%;  Score 48.5;  DB 2;  Length 608;
  Best Local Similarity   33.3%;  Pred. No. 53;
  Matches    9;  Conservative    9;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          1 PIN-GTNSLTKKVFGLKKDGDITKKDD 26
              |:| || ::    ||: :  :::|:||
Db        554 PVNKGTTTVDGAGFGIDRPAELSKEDD 580


RESULT 11
JN0635
neural cell adhesion molecule 2 precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: JN0635
R;Tonissen, K.F.; Krieg, P.A.
Gene 127, 243-247, 1993
A;Title: Two neural-cell adhesion molecule(NCAM)-encoding genes in Xenopus laevis are exp
A;Reference number: JN0635; MUID:93273239; PMID:7684721
A;Accession: JN0635
A;Molecule type: mRNA
A;Residues: 1-1092 <TON>
A;Cross-references: UNIPROT:P36335; UNIPARC:UPI000012FDC7; GB:M76710; NID:g214611; PIDN:A
C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mole
C;Genetics:
A;Gene: NCAM2
C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immun
C;Keywords: alternative splicing; cell adhesion; duplication; heparin binding; sialoglyco
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1092/Product: neural cell adhesion molecule 2 #status predicted <NCA>
F;20-705/Domain: extracellular #status predicted <EXT>
F;34-95/Domain: immunoglobulin homology <IMM1>
F;129-188/Domain: immunoglobulin homology <IMM2>
F;149-153/Region: heparin binding #status predicted
F;158-162/Region: heparin binding #status predicted
F;317-381/Domain: immunoglobulin homology <IMM3>
F;413-475/Domain: immunoglobulin homology <IMM4>
F;512-589/Domain: fibronectin type III repeat homology <FN3A>
F;619-680/Domain: fibronectin type III repeat homology <FN3B>
F;706-723/Domain: transmembrane #status predicted <TMM>
F;724-1092/Domain: intracellular #status predicted <INT>
F;41-93,136-186,232-282,323-379,420-473/Disulfide bonds: #status predicted
F;219,310,341,417,443,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             35.7%;  Score 48.5;  DB 1;  Length 1092;
  Best Local Similarity   41.9%;  Pred. No. 98;
  Matches   13;  Conservative    3;  Mismatches   10;  Indels    5;  Gaps    1;

Qy          1 PINGTNSLTKKVFGLKKDG-----DITKKDD 26
              |: |  |  | |  | :||     || |:||
Db        591 PVKGEPSAPKLVGHLSEDGNSIKVDIIKQDD 621
```

```
RESULT 12
E71156
endopeptidase La homolog (EC 3.4.21.-) - Pyrococcus horikoshii
N;Alternate names: ATP-dependent proteinase lon homolog
C;Species: Pyrococcus horikoshii
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 16-Aug-2004
C;Accession: E71156
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71156
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1127 <KAW>
A;Cross-references: UNIPROT:O58221; UNIPARC:UPI000012E7ED; GB:AP000002; NID:g3236129; PI
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0452
C;Keywords: ATP; hydrolase; nucleotide binding; P-loop; serine proteinase
F;19-98/Domain: Methanococcus endopeptidase La homolog P-loop-containing homology <MLAP>
F;59-66/Region: nucleotide-binding motif A (P-loop)
F;708-713/Region: nucleotide-binding motif B
F;990/Active site: Ser #status predicted

  Query Match             35.7%;  Score 48.5;  DB 1;  Length 1127;
  Best Local Similarity   57.9%;  Pred. No. 1e+02;
  Matches   11;  Conservative    4;  Mismatches    1;  Indels    3;  Gaps    1;

Qy          11 KVF---GLKKDGDITKKDD 26
               |||   |||: |:||:||:
Db         305 KVFTSEGLKEAGEITEKDE 323


RESULT 13
T29928
hypothetical protein F29B9.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29928
R;Goela, D.; Gattung, S.
submitted to the EMBL Data Library, September 1996
A;Description: The sequence of C. elegans cosmid F29B9.
A;Reference number: Z20710
A;Accession: T29928
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-155 <GOE>
A;Cross-references: UNIPROT:Q9GYI6; UNIPARC:UPI000007C274; EMBL:U70849; PIDN:AAB09113.1;
A;Experimental source: strain Bristol N2; clone F29B9
C;Genetics:
A;Gene: CESP:F29B9.7
A;Map position: 4
A;Introns: 43/2; 69/2; 97/2; 135/2
C;Superfamily: Caenorhabditis elegans hypothetical protein F29B9.7

  Query Match             35.3%;  Score 48;  DB 2;  Length 155;
  Best Local Similarity   43.5%;  Pred. No. 15;
  Matches   10;  Conservative    6;  Mismatches    3;  Indels    4;  Gaps    1;

Qy           4 GTNSLTKKVFGLKKDGDITKKDD 26
               |: | ||:    :|:|:: ||||
Db         137 GSRSPTKE----RKEGEVLKKDD 155


RESULT 14
H83707
bacitracin resistance protein (undecaprenol kinase) BH0464 [imported] - Bacillus halodur
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83707
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and g
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83707
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-274 <STO>
A;Cross-references: UNIPROT:Q9KFL5; UNIPARC:UPI0000137CCB; GB:AP001508; GB:BA000004; NID:
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0464
C;Superfamily: Escherichia coli bacitracin resistance protein bacA

  Query Match             35.3%;  Score 48;  DB 2;  Length 274;
  Best Local Similarity   55.6%;  Pred. No. 27;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy           8 LTKKVFGLKKDGDITKKD 25
               | : : | |:|||| ||
Db         139 LIRHLNGRKQDGDITLKD 156


RESULT 15
F75169
serine hydroxymethyltransferase (glya) PAB2018 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F75169
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
A;Reference number: A75001
A;Accession: F75169
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <KAW>
A;Cross-references: UNIPROT:Q9V1B2; UNIPARC:UPI000003468C; GB:AJ248284; GB:AL096836; NID:
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB2018
C;Superfamily: serine/glycine hydroxymethyltransferase

  Query Match             35.3%;  Score 48;  DB 2;  Length 427;
  Best Local Similarity   50.0%;  Pred. No. 43;
  Matches   10;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy           1 PINGTNSLTKKVFGLKKDGD 20
               ||:|||:     ||| : ||
Db          90 PISGTNANQAVFFGLGQPGD 109


Search completed: January 24, 2006, 11:47:48
Job time : 5.07843 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:28:32 ; Search time 190.327 Seconds
(without alignments)
646.394 Million cell updates/sec

Title: US-10-680-349-42
Perfect score: 1462
Sequence: 1 MNYKKILVRSALISLMSILP.....ASVTLDVGYFGGEIGMRFTF 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1462	100.0	280	5	AAU96116	Aau96116 Ehrlichia
2	1462	100.0	280	5	ABG77958	Abg77958 Ehrlichia
3	1462	100.0	280	6	ADA09781	Ada09781 E. canis
4	1462	100.0	280	9	ADW04274	Adw04274 Ehrlichia
5	1202.5	82.3	283	2	AAY06944	Aay06944 E. chafee
6	1202.5	82.3	283	5	AAU96106	Aau96106 Ehrlichia
7	1202.5	82.3	283	5	AAU73413	Aau73413 Ehrlichia
8	1202.5	82.3	283	5	ABG77936	Abg77936 Ehrlichia
9	1202.5	82.3	283	6	ADA09737	Ada09737 E. chaffe
10	1202.5	82.3	283	9	ADW04230	Adw04230 Erlichia
11	715	48.9	165	2	AAY06970	Aay06970 E. canis
12	644.5	44.1	281	2	AAY06943	Aay06943 E. chafee
13	644.5	44.1	281	5	AAU96105	Aau96105 Ehrlichia
14	644.5	44.1	281	5	AAU73418	Aau73418 Ehrlichia
15	644.5	44.1	281	5	ABG77935	Abg77935 Ehrlichia
16	642.5	43.9	281	6	ADA09735	Ada09735 E. chaffe
17	642.5	43.9	281	9	ADW04228	Adw04228 Erlichia
18	642.5	43.9	281	9	ADW04293	Adw04293 Erlichia
19	629.5	43.1	276	2	AAW51095	Aaw51095 Ehrlichia
20	629.5	43.1	276	3	AAB36189	Aab36189 Ehrlichia
21	629.5	43.1	276	4	AAU04199	Aau04199 Variable
22	622	42.5	286	2	AAW51092	Aaw51092 Ehrlichia
23	621	42.5	288	2	AAY06959	Aay06959 E. canis
24	621	42.5	288	5	ABG77950	Abg77950 Ehrlichia
25	621	42.5	288	6	ADA09765	Ada09765 E. canis
26	621	42.5	288	9	ADW04258	Adw04258 Ehrlichia
27	620	42.4	286	2	AAY06946	Aay06946 E. chafee
28	620	42.4	286	3	AAB36186	Aab36186 Ehrlichia
29	620	42.4	286	4	AAU04196	Aau04196 Variable
30	620	42.4	286	5	AAU96108	Aau96108 Ehrlichia
31	620	42.4	286	5	AAU73415	Aau73415 Ehrlichia
32	620	42.4	286	5	ABG77938	Abg77938 Ehrlichia
33	620	42.4	286	6	ADA09741	Ada09741 E. chaffe
34	620	42.4	286	9	ADW04234	Adw04234 Erlichia
35	606	41.5	280	5	ABG77953	Abg77953 Ehrlichia
36	605	41.4	280	2	AAY06948	Aay06948 E. chafee
37	605	41.4	280	5	AAU96110	Aau96110 Ehrlichia
38	605	41.4	280	5	ABG77940	Abg77940 Ehrlichia
39	605	41.4	280	6	ADA09745	Ada09745 E. chaffe
40	605	41.4	280	9	ADW04238	Adw04238 Erlichia
41	603	41.2	280	2	AAW51094	Aaw51094 Ehrlichia
42	603	41.2	280	3	AAB36188	Aab36188 Ehrlichia
43	603	41.2	280	4	AAU04198	Aau04198 Variable
44	603	41.2	280	5	AAU73417	Aau73417 Ehrlichia
45	601	41.1	280	2	AAY06962	Aay06962 E. canis

ALIGNMENTS

RESULT 1
AAU96116
ID AAU96116 standard; protein; 280 AA.
XX
AC AAU96116;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ehrlichia canis p28-2.
XX
KW Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.
XX
OS Ehrlichia canis.
XX
PN WO200222782-A2.
XX
PD 21-MAR-2002.
XX
PF 12-SEP-2001; 2001WO-US028759.
XX
PR 12-SEP-2000; 2000US-00660587.
XX
PA (RERE-) RES DEV FOUND.
XX
PI Walker DH, Yu X, Mcbride JW;
XX
DR WPI; 2002-351882/38.
DR N-PSDB; ABK68876.
XX
PT New recombinant homologous 28 kilodalton immunodominant protein from
PT Ehrlichia canis, useful for treating Ehrlichia canis infections.
XX
PS Claim 16; Fig 14; 106pp; English.
XX
CC The invention relates to a recombinant homologous 28 kDa immunodominant
CC protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably
CC dispersed in a pharmaceutically acceptable carrier, is useful for
CC inhibiting E. canis infection in a subject. (I) is useful in the
CC development of vaccines and serodiagnostics that are particularly
CC effective for disease prevention and serodiagnosis. AAU96100-AAU96118
CC represent the 28-kDa antigen amino acid sequences of the invention
XX
SQ Sequence 280 AA;

Query Match 100.0%; Score 1462; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.1e-138;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 2
ABG77958
ID ABG77958 standard; protein; 280 AA.
XX
AC ABG77958;
XX
DT 15-NOV-2002 (first entry)
XX
DE Ehrlichia canis outer membrane protein (P30F) #9.
XX
KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.
XX
OS Ehrlichia canis.
XX
PN US2002120115-A1.
XX
PD 29-AUG-2002.
XX
PF 28-JAN-2002; 2002US-00059964.
XX
PR 19-MAY-1999; 99US-00314701.
XX
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2002-618954/66.
DR N-PSDB; ABS63299.
XX
PT Isolated polynucleotide encoding an outer membrane protein of E.canis or
PT E.chaffeensis used in the diagnosis of infection.
XX
PS Claim 10; Fig 30B; 49pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding an outer
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
CC in the diagnosis of infection. An infection such as human ehrlichiosis or
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC patient, providing a polypeptide or mixture of polypeptides, contacting
CC the sample with the polypeptide and assaying for the formation of a
CC complex between antibodies in the serum sample and the polypeptide, where
CC formation of a complex is indicative of infection with E. chaffeensis.
CC This sequence represents an Ehrlichia outer membrane protein of the
CC invention
XX
SQ Sequence 280 AA;

Query Match 100.0%; Score 1462; DB 5; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.1e-138;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
```

RESULT 3
ADA09781
ID ADA09781 standard; protein; 280 AA.
XX
AC ADA09781;
XX
DT 06-NOV-2003 (first entry)
XX
DE E. canis outer membrane protein P30-10.
XX
KW outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;
KW Rocky Mountian spotted fever; canine ehrlichiosis; antigen.
XX
OS Ehrlichia canis.
XX
PN US6544517-B1.
XX
PD 08-APR-2003.
XX
PF 19-MAY-1999; 99US-00314701.
XX
PR 18-SEP-1998; 98US-0100843P.
XX
PA (OHIS) UNIV OHIO STATE RES FOUND.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2003-553952/52.
DR N-PSDB; ADA09780.
XX
PT New isolated polynucleotide encoding outer membrane protein P30 of
PT Ehrlichia canis or its variant or fragment, useful for producing
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for
PT diagnosing and treating ehrlichiosis.
XX
PS Disclosure; Fig 30; 105pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a variant of
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer
CC membrane protein of E. canis , or an antigenic fragment of the E. canis
CC P30 protein, or comprising a sequence which is the complement of
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.
CC canis and E. chaffeensis outer membrane proteins and their encoding
CC nucleic acids. The polynucleotides are useful for producing E. canis or
CC E. chaffeensis outer membrane protein, for designing hybridisation probes
CC for isolating and identifying cDNA and genomic clones encoding the OMP)
CC or its allelic forms, for designing primers for PCR. The polypeptides
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present
CC sequence represents an E. canis outer membrane protein.
XX

```
SQ   Sequence 280 AA;

  Query Match            100.0%;  Score 1462;  DB 6;  Length 280;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-138;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 4
ADW04274
ID   ADW04274 standard; protein; 280 AA.
XX
AC   ADW04274;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Ehrlichia canis outer membrane protein (P30F), P30-10.
XX
KW   DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;
KW   infection; vaccine.
XX
OS   Ehrlichia canis.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .25
FT                   /label= Signal_peptide
FT   Protein         26. .280
FT                   /note= "Erlichia canis mature OMP protein"
XX
PN   US2004265334-A1.
XX
PD   30-DEC-2004.
XX
PF   29-JUL-2004; 2004US-00901774.
XX
PR   18-SEP-1998;   98US-0100843P.
PR   19-MAY-1999;   99US-00314701.
PR   28-JAN-2002; 2002US-00059964.
XX
PA   (RIKI/) RIKIHISA Y.
PA   (OHAS/) OHASHI N.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 2005-064871/07.
DR   N-PSDB; ADW04273.
XX
PT   New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia
PT   canis or Ehrlichia chaffeensis, useful in preparing a composition for
PT   diagnosing or preventing E. canis or E. chaffeensis infection.
XX
PS   Claim 10; SEQ ID NO 48; 122pp; English.
XX
CC   The invention relates to nucleic acid sequences encoding outer membrane
CC   proteins (OMP) of Erlichia chaffeensis (designated as OMP proteins) and
CC   Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is
CC   useful in preparing a composition for diagnosing, treating or preventing
CC   an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present
CC   sequence is the Erlichia canis P30F protein.
XX
SQ   Sequence 280 AA;

  Query Match            100.0%;  Score 1462;  DB 9;  Length 280;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-138;
  Matches  280;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCYDIT 180

Qy        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVIGNK 240

Qy        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              ||||||||||||||||||||||||||||||||||||||||
Db        241 FEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280


RESULT 5
AAY06944
ID   AAY06944 standard; protein; 283 AA.
XX
AC   AAY06944;
XX
DT   27-AUG-2003  (revised)
DT   05-JUL-1999  (first entry)
XX
DE   E. chafeensis OMP-1B protein.
XX
KW   Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW   detection; dog.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO9913720-A1.
XX
PD   25-MAR-1999.
XX
PF   18-SEP-1998;   98WO-US019600.
XX
PR   19-SEP-1997;   97US-0059353P.
XX
PA   (OHIS ) UNIV OHIO STATE.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 1999-254290/21.
DR   N-PSDB; AAX34744.
XX
PT   Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
PT   canis.
XX
PS   Claim 12; Fig 4B; 55pp; English.
XX
CC   The invention provides isolated outer membrane proteins (OMP) from
CC   Ehrlichia chafeensis and E. canis. The E. chaffeensis proteins form part
CC   of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in
CC   AAY06943-958. The E. canis proteins form part of the P30 family and
```

```
CC   consist of proteins shown in AAY06959-970. The proteins and genes are
CC   used to detect E. chaffeensis in patients and E. canis in dogs. (Updated
CC   on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 283 AA;

  Query Match             82.3%;  Score 1202.5;  DB 2;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 3e-112;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 6
AAU96106
ID   AAU96106 standard; protein; 283 AA.
XX
AC   AAU96106;
XX
DT   07-AUG-2003  (revised)
DT   02-JUL-2002  (first entry)
XX
DE   Ehrlichia chafeensis OMP-1B.
XX
KW   Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO200222782-A2.
XX
PD   21-MAR-2002.
XX
PF   12-SEP-2001; 2001WO-US028759.
XX
PR   12-SEP-2000; 2000US-00660587.
XX
PA   (RERE-) RES DEV FOUND.
XX
PI   Walker DH,  Yu X,  Mcbride JW;
XX
DR   WPI; 2002-351882/38.
XX
PT   New recombinant homologous 28 kilodalton immunodominant protein from
PT   Ehrlichia canis, useful for treating Ehrlichia canis infections.
XX
PS   Example 3; Fig 3; 106pp; English.
XX
CC   The invention relates to a recombinant homologous 28 kDa immunodominant
CC   protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably
CC   dispersed in a pharmaceutically acceptable carrier, is useful for
CC   inhibiting E. canis infection in a subject. (I) is useful in the
CC   development of vaccines and serodiagnostics that are particularly
CC   effective for disease prevention and serodiagnosis. AAU96100-AAU96118
CC   represent the 28-kDa antigen amino acid sequences of the invention.
CC   (Updated on 07-AUG-2003 to correct OS field.)
XX
SQ   Sequence 283 AA;

  Query Match             82.3%;  Score 1202.5;  DB 5;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 3e-112;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:||||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 7
AAU73413
ID   AAU73413 standard; protein; 283 AA.
XX
AC   AAU73413;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Ehrlichia chaffeensis outer membrane protein P28-14.
XX
KW   Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO200183699-A2.
XX
PD   08-NOV-2001.
XX
PF   01-MAY-2001; 2001WO-US013997.
XX
PR   01-MAY-2000; 2000US-0201035P.
XX
PA   (RERE-) RES DEV FOUND.
XX
PI   Walker DH,  Yu X;
XX
DR   WPI; 2002-066527/09.
XX
PT   Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28
PT   useful as a vaccine against Ehrlichia chaffeensis.
XX
PS   Disclosure; Fig 2; 97pp; English.
XX
CC   The invention relates to isolated and purified 28-kDa outer membrane
CC   proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are
CC   encoded by a 28kDa outer membrane protein multigene family. P28 proteins
CC   are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful
CC   for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia
CC   chaffeensis P28 outer membrane proteins of the invention
XX
SQ   Sequence 283 AA;

  Query Match             82.3%;  Score 1202.5;  DB 5;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 3e-112;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;
```

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283
```

RESULT 8
ABG77936
ID ABG77936 standard; protein; 283 AA.
XX
AC ABG77936;
XX
DT 15-NOV-2002 (first entry)
XX
DE Ehrlichia chaffeensis outer membrane protein (OMP) #2.
XX
KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.
XX
OS Ehrlichia chaffeensis.
XX
PN US2002120115-A1.
XX
PD 29-AUG-2002.
XX
PF 28-JAN-2002; 2002US-00059964.
XX
PR 19-MAY-1999; 99US-00314701.
XX
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2002-618954/66.
DR N-PSDB; ABS63277.
XX
PT Isolated polynucleotide encoding an outer membrane protein of E.canis or
PT E.chaffeensis used in the diagnosis of infection.
XX
PS Disclosure; Fig 4B; 49pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding an outer
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
CC in the diagnosis of infection. An infection such as human ehrlichiosis or
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC patient, providing a polypeptide or mixture of polypeptides, contacting
CC the sample with the polypeptide and assaying for the formation of a
CC complex between antibodies in the serum sample and the polypeptide, where
CC formation of a complex is indicative of infection with E. chaffeensis.
CC This sequence represents an Ehrlichia outer membrane protein of the
CC invention
XX
SQ Sequence 283 AA;

Query Match 82.3%; Score 1202.5; DB 5; Length 283;
Best Local Similarity 79.5%; Pred. No. 3e-112;
Matches 225; Conservative 26; Mismatches 29; Indels 3; Gaps 1;

```
Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:||||| :||| ||||:||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283
```

RESULT 9
ADA09737
ID ADA09737 standard; protein; 283 AA.
XX
AC ADA09737;
XX
DT 06-NOV-2003 (first entry)
XX
DE E. chaffeensis outer membrane protein OMP-1B.
XX
KW outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;
KW Rocky Mountian spotted fever; canine ehrlichiosis; antigen.
XX
OS Ehrlichia chaffeensis.
XX
PN US6544517-B1.
XX
PD 08-APR-2003.
XX
PF 19-MAY-1999; 99US-00314701.
XX
PR 18-SEP-1998; 98US-0100843P.
XX
PA (OHIS) UNIV OHIO STATE RES FOUND.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2003-553952/52.
DR N-PSDB; ADA09736.
XX
PT New isolated polynucleotide encoding outer membrane protein P30 of
PT Ehrlichia canis or its variant or fragment, useful for producing
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for
PT diagnosing and treating ehrlichiosis.
XX
PS Disclosure; Fig 4; 105pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a variant of
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer
CC membrane protein of E. canis , or an antigenic fragment of the E. canis
CC P30 protein, or comprising a sequence which is the complement of
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.
CC canis and E. chaffeensis outer membrane proteins and their encoding
CC nucleic acids. The polynucleotides are useful for producing E. canis or
CC E. chaffeensis outer membrane protein, for designing hybridisation probes
CC for isolating and identifying cDNA and genomic clones encoding the OMP)
CC or its allelic forms, for designing primers for PCR. The polypeptides
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present
CC sequence represents an E. chaffeensis outer membrane protein.
XX

```
SQ   Sequence 283 AA;

  Query Match             82.3%;  Score 1202.5;  DB 6;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 3e-112;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:|||||  :||| ||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 10
ADW04230
ID   ADW04230 standard; protein; 283 AA.
XX
AC   ADW04230;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Erlichia chaffeensis outer membrane protein (OMP), OMP-1B.
XX
KW   DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;
KW   infection; vaccine.
XX
OS   Ehrlichia chaffeensis.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .25
FT                   /label= Signal_peptide
FT   Protein         26. .283
FT                   /note= "Erlichia chaffeensis mature OMP protein"
FT   Region          26. .41
FT                   /note= "Semivariable region (SV)"
FT   Region          82. .94
FT                   /note= "Hypervariable region (HV1)"
FT   Region          145. .163
FT                   /note= "Hypervariable region (HV2)"
FT   Region          248. .272
FT                   /note= "Hypervariable region (HV3)"
XX
PN   US2004265334-A1.
XX
PD   30-DEC-2004.
XX
PF   29-JUL-2004; 2004US-00901774.
XX
PR   18-SEP-1998;   98US-0100843P.
PR   19-MAY-1999;   99US-00314701.
PR   28-JAN-2002; 2002US-00059964.
XX
PA   (RIKI/) RIKIHISA Y.
PA   (OHAS/) OHASHI N.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 2005-064871/07.
DR   N-PSDB; ADW04229.
XX
PT   New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia
PT   canis or Ehrlichia chaffeensis, useful in preparing a composition for
PT   diagnosing or preventing E. canis or E. chaffeensis infection.
XX
PS   Disclosure; SEQ ID NO 4; 122pp; English.
XX
CC   The invention relates to nucleic acid sequences encoding outer membrane
CC   proteins (OMP) of Erlichia chaffeensis (designated as OMP proteins) and
CC   Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is
CC   useful in preparing a composition for diagnosing, treating or preventing
CC   an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present
CC   sequence is the Erlichia chaffeensis OMP protein.
XX
SQ   Sequence 283 AA;

  Query Match             82.3%;  Score 1202.5;  DB 9;  Length 283;
  Best Local Similarity   79.5%;  Pred. No. 3e-112;
  Matches  225;  Conservative   26;  Mismatches   29;  Indels    3;  Gaps    1;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSR---TNDNKEGFYISAKYNPSISHFRKFSA 57
              |||||| | |||||||||||||||||||| |     ||::|||||| ||||||||||||||
Db          1 MNYKKIFVSSALISLMSILPYQSFADPVTSNDTGINDSREGFYISVKYNPSISHFRKFSA 60

Qy         58 EETPINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRI 117
              || ||||  |:||||||||||||| :  :| |  | ::|||||||||||||||:||||||
Db         61 EEAPINGNTSITKKVFGLKKDGDIAQSANFNRTDPALEFQNNLISGFSGSIGYAMDGPRI 120

Qy        118 ELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||||||:|: || |||||::|:|||:| |||:||: |::||||||:|||||||||||||
Db        121 ELEAAYQKFDAKNPDNNDTNSGDYYKYFGLSREDAIADKKYVVLKNEGITFMSLMVNTCY 180

Qy        178 DITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIGGYYHGVI 237
              ||||||| |:||||||:|||||  :||| ||||:|||||||||||||||||||||||||||
Db        181 DITAEGVPFIPYACAGVGADLINVFKDFNLKFSYQGKIGISYPITPEVSAFIGGYYHGVI 240

Qy        238 GNKFEKIPVITPVVLNDAPQTTSASVTLDVGYFGGEIGMRFTF 280
              || | ||||||||||  ||||||| ||:| ||||||:|:||||
Db        241 GNNFNKIPVITPVVLEGAPQTTSALVTIDTGYFGGEVGVRFTF 283


RESULT 11
AAY06970
ID   AAY06970 standard; protein; 165 AA.
XX
AC   AAY06970;
XX
DT   05-JUL-1999  (first entry)
XX
DE   E. canis P30-10protein.
XX
KW   Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW   detection; dog.
XX
OS   Ehrlichia canis.
XX
PN   WO9913720-A1.
XX
PD   25-MAR-1999.
XX
PF   18-SEP-1998;   98WO-US019600.
XX
PR   19-SEP-1997;   97US-0059353P.
XX
PA   (OHIS ) UNIV OHIO STATE.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 1999-254290/21.
DR   N-PSDB; AAX34770.
XX
PT   Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
PT   canis.
```

```
XX
PS   Disclosure; Fig 30B; 55pp; English.
XX
CC   The invention provides isolated outer membrane proteins (OMP) from
CC   Ehrlichia chafeensis and E. canis. The E. chaffeensis proteins form part
CC   of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in
CC   AAY06943-958. The E. canis proteins form part of the P30 family and
CC   consist of proteins shown in AAY06959-970. The proteins and genes are
CC   used to detect E. chaffeensis in patients and E. canis in dogs
XX
SQ   Sequence 165 AA;

  Query Match             48.9%;  Score 715;  DB 2;  Length 165;
  Best Local Similarity   81.9%;  Pred. No. 1.8e-63;
  Matches  145;  Conservative    4;  Mismatches   16;  Indels   12;  Gaps    2;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60

Qy         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDGPRIELE 120

Qy        121 AAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAMEDQQYVVLKNDGITFMSLMVNTCY 177
              |||     |: |||||||||||||||           |:|     |: |  : '| |
Db        121 AAYHNLIQKH-DNNDTDNGEYYKHFA-----------YLVKMPWKISHMLFLKMTAY 165


RESULT 12
AAY06943
ID   AAY06943 standard; protein; 281 AA.
XX
AC   AAY06943;
XX
DT   27-AUG-2003  (revised)
DT   05-JUL-1999  (first entry)
XX
DE   E. chafeensis OMP-1 protein.
XX
KW   Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW   detection; dog.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO9913720-A1.
XX
PD   25-MAR-1999.
XX
PF   18-SEP-1998;   98WO-US019600.
XX
PR   19-SEP-1997;   97US-0059353P.
XX
PA   (OHIS ) UNIV OHIO STATE.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 1999-254290/21.
DR   N-PSDB; AAX34743.
XX
PT   Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
PT   canis.
XX
PS   Disclosure; Fig 3B; 55pp; English.
XX
CC   The invention provides isolated outer membrane proteins (OMP) from
CC   Ehrlichia chafeensis and E. canis. The E. chaffeensis proteins form part
CC   of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in
CC   AAY06943-958. The E. canis proteins form part of the P30 family and
CC   consist of proteins shown in AAY06959-970. The proteins and genes are
CC   used to detect E. chaffeensis in patients and E. canis in dogs. (Updated
CC   on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 281 AA;

  Query Match             44.1%;  Score 644.5;  DB 2;  Length 281;
  Best Local Similarity   48.1%;  Pred. No. 5e-56;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : | ||||||:: |       :  | ||       ::||   ||:|:|||||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||      | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
              |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              |::| ||||:|  || | |     | :    | | ||| :|| |:| || |
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFVF 281


RESULT 13
AAU96105
ID   AAU96105 standard; protein; 281 AA.
XX
AC   AAU96105;
XX
DT   07-AUG-2003  (revised)
DT   02-JUL-2002  (first entry)
XX
DE   Ehrlichia chafeensis P28.
XX
KW   Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO200222782-A2.
XX
PD   21-MAR-2002.
XX
PF   12-SEP-2001; 2001WO-US028759.
XX
PR   12-SEP-2000; 2000US-00660587.
XX
PA   (RERE-) RES DEV FOUND.
XX
PI   Walker DH,  Yu X,  Mcbride JW;
XX
DR   WPI; 2002-351882/38.
XX
PT   New recombinant homologous 28 kilodalton immunodominant protein from
PT   Ehrlichia canis, useful for treating Ehrlichia canis infections.
XX
PS   Example 3; Fig 3; 106pp; English.
XX
CC   The invention relates to a recombinant homologous 28 kDa immunodominant
CC   protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably
CC   dispersed in a pharmaceutically acceptable carrier, is useful for
CC   inhibiting E. canis infection in a subject. (I) is useful in the
CC   development of vaccines and serodiagnostics that are particularly
CC   effective for disease prevention and serodiagnosis. AAU96100-AAU96118
CC   represent the 28-kDa antigen amino acid sequences of the invention.
CC   (Updated on 07-AUG-2003 to correct OS field.)
XX
SQ   Sequence 281 AA;

  Query Match             44.1%;  Score 644.5;  DB 5;  Length 281;
```

```
  Best Local Similarity   48.1%;  Pred. No. 5e-56;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:: |        :  | ||       ::||   ||:|:|||||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||      | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
              |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              |::| ||||:|  || | |     | :    | | ||| :|| |:| || |
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFAF 281


RESULT 14
AAU73418
ID   AAU73418 standard; protein; 281 AA.
XX
AC   AAU73418;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Ehrlichia chaffeensis outer membrane protein P28-19.
XX
KW   Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO200183699-A2.
XX
PD   08-NOV-2001.
XX
PF   01-MAY-2001; 2001WO-US013997.
XX
PR   01-MAY-2000; 2000US-0201035P.
XX
PA   (RERE-) RES DEV FOUND.
XX
PI   Walker DH,  Yu X;
XX
DR   WPI; 2002-066527/09.
XX
PT   Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28
PT   useful as a vaccine against Ehrlichia chaffeensis.
XX
PS   Disclosure; Fig 2; 97pp; English.
XX
CC   The invention relates to isolated and purified 28-kDa outer membrane
CC   proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are
CC   encoded by a 28kDa outer membrane protein multigene family. P28 proteins
CC   are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful
CC   for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia
CC   chaffeensis P28 outer membrane proteins of the invention
XX
SQ   Sequence 281 AA;

  Query Match           44.1%;  Score 644.5;  DB 5;  Length 281;
  Best Local Similarity   48.1%;  Pred. No. 5e-56;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56

Qy         61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
                 : |  |||||:: |        :  | ||       ::||   ||:|:|||||||
Db         57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy        115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
              |||||| :|: |: ||  ||      | ::: |||   | :       :| |||:|:  :|
Db        113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy        171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
              |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db        171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy        231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
              |::| ||||:|  || | |     | :    | | ||| :|| |:| || |
Db        231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFAF 281


RESULT 15
ABG77935
ID   ABG77935 standard; protein; 281 AA.
XX
AC   ABG77935;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Ehrlichia chaffeensis outer membrane protein (OMP) #1.
XX
KW   Outer membrane protein; OMP; P30F; ehrlichiosis; infection.
XX
OS   Ehrlichia chaffeensis.
XX
PN   US2002120115-A1.
XX
PD   29-AUG-2002.
XX
PF   28-JAN-2002; 2002US-00059964.
XX
PR   19-MAY-1999;   99US-00314701.
XX
PA   (RIKI/) RIKIHISA Y.
PA   (OHAS/) OHASHI N.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 2002-618954/66.
DR   N-PSDB; ABS63276.
XX
PT   Isolated polynucleotide encoding an outer membrane protein of E.canis or
PT   E.chaffeensis used in the diagnosis of infection.
XX
PS   Claim 14; Fig 3B; 49pp; English.
XX
CC   The invention relates to an isolated polynucleotide encoding an outer
CC   membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
CC   in the diagnosis of infection. An infection such as human ehrlichiosis or
CC   canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC   patient, providing a polypeptide or mixture of polypeptides, contacting
CC   the sample with the polypeptide and assaying for the formation of a
CC   complex between antibodies in the serum sample and the polypeptide, where
CC   formation of a complex is indicative of infection with E. chaffeensis.
CC   This sequence represents an Ehrlichia outer membrane protein of the
CC   invention
XX
SQ   Sequence 281 AA;

  Query Match           44.1%;  Score 644.5;  DB 5;  Length 281;
  Best Local Similarity   48.1%;  Pred. No. 5e-56;
  Matches  140;  Conservative   41;  Mismatches   89;  Indels   21;  Gaps    6;

Qy          1 MNYKKILVRSALISLMSILPYQSFADPVGSRTNDNKEGFYISAKYNPSISHFRKFSAEET 60
              |||||: : ||||||:| ||  ||:|| ||  | |   |||| || || |||  |||:|
Db          1 MNYKKVFITSALISLISSLPGVSFSDPAGSGINGN---FYISGKYMPSASHFGVFSAKE- 56
```

```
Qy          61 PINGTNSLTKKVFGLKKDGD------ITKKDDFTRVAPGIDFQNNLISGFSGSIGYSMDG 114
               :  |  |||||:: |       :  | ||       ::||   ||:|:|||||||
Db          57 ----ERNTTVGVFGLKQNWDGSAISNSSPNDVFTVSNYSFKYENNPFLGFAGAIGYSMDG 112

Qy         115 PRIELEAAYQQFNPKNTDNNDTDNGEYYKHFALSRKDAME----DQQYVVLKNDGITFMS 170
               |||||| :|: |: ||  ||      | ::: |||   | :       :| |||:|:  :|
Db         113 PRIELEVSYETFDVKNQGNN--YKNEAHRYCALSHNSAADMSSASNNFVFLKNEGLLDIS 170

Qy         171 LMVNTCYDITAEGVSFVPYACAGIGADLITIFKDLNLKFAYQGKIGISYPITPEVSAFIG 230
                |:| |||:  ||: | || ||||| ||:::|:  | | :||||:|:|| |:|| | |||
Db         171 FMLNACYDVVGEGIPFSPYICAGIGTDLVSMFEATNPKISYQGKLGLSYSISPEASVFIG 230

Qy         231 GYYHGVIGNKFEKIPVITPVVLNDAPQ-TTSASVTLDVGYFGGEIGMRFTF 280
               |::| ||||:|  || | |     | :    | | ||| :|| |:| || |
Db         231 GHFHKVIGNEFRDIPTIIPTGSTLAGKGNYPAIVILDVCHFGIELGGRFAF 281
```

Search completed: January 24, 2006, 11:42:45
Job time : 191.327 secs

This page blank (uspto)

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 24, 2006, 11:28:32 ; Search time 17.6732 Seconds
(without alignments)
646.394 Million cell updates/sec

Title: US-10-680-349-42_COPY_61_86
Perfect score: 136
Sequence: 1 PINGTNSLTKKVFGLKKDGDITKKDD 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	136	100.0	165	2	AAY06970	Aay06970 E. canis
2	136	100.0	280	5	AAU96116	Aau96116 Ehrlichia
3	136	100.0	280	5	ABG77958	Abg77958 Ehrlichia
4	136	100.0	280	6	ADA09781	Ada09781 E. canis
5	136	100.0	280	9	ADW04274	Adw04274 Ehrlichia
6	97	71.3	283	2	AAY06944	Aay06944 E. chafee
7	97	71.3	283	5	AAU96106	Aau96106 Ehrlichia
8	97	71.3	283	5	AAU73413	Aau73413 Ehrlichia
9	97	71.3	283	5	ABG77936	Abg77936 Ehrlichia
10	97	71.3	283	6	ADA09737	Ada09737 E. chaffe
11	97	71.3	283	9	ADW04230	Adw04230 Erlichia
12	52.5	38.6	339	8	ADN21449	Adn21449 Bacterial
13	52	38.2	377	6	ABU29155	Abu29155 Protein e
14	52	38.2	378	7	ADH88024	Adh88024 Enterococ
15	52	38.2	481	6	ABU20708	Abu20708 Protein e
16	51	37.5	525	8	ADS44601	Ads44601 Bacterial
17	50.5	37.1	873	8	ADS22299	Ads22299 Bacterial
18	50	36.8	283	9	ADW04295	Adw04295 Cowdria r
19	50	36.8	284	5	AAU96111	Aau96111 Cowdria r
20	50	36.8	302	6	ABP79144	Abp79144 N. gonorr
21	50	36.8	302	6	ABP76775	Abp76775 N. gonorr
22	49.5	36.4	521	7	ABO23592	Abo23592 Rickettsi
23	49	36.0	235	9	ADW17535	Adw17535 Pinus rad
24	49	36.0	345	7	ADC00818	Adc00818 Enterohae
25	49	36.0	472	8	ADN46936	Adn46936 Thermococ
26	49	36.0	511	5	ABP65734	Abp65734 Bifidobac
27	49	36.0	577	5	ABP27646	Abp27646 Streptoco
28	49	36.0	577	6	ABU46702	Abu46702 Protein e
29	49	36.0	629	7	ADD68790	Add68790 Streptoco
30	49	36.0	629	8	ADJ61994	Adj61994 Group A s
31	49	36.0	1992	2	AAW04505	Aaw04505 Moraxella
32	49	36.0	1992	4	AAB69137	Aab69137 M. catarr
33	49	36.0	1992	4	AAB69133	Aab69133 M. catarr
34	49	36.0	2047	4	AAB69134	Aab69134 M. catarr
35	48.5	35.7	416	2	AAR74151	Aar74151 Mango cla
36	48.5	35.7	416	2	AAW44331	Aaw44331 Class II
37	48.5	35.7	476	4	ABG99093	Abg99093 Pyrococcu
38	48.5	35.7	476	8	ADL23827	Adl23827 Pyrococcu
39	48.5	35.7	557	4	AAU20133	Aau20133 Human DNA
40	48.5	35.7	557	5	ABG91382	Abg91382 Novel hum
41	48.5	35.7	585	8	ADG32050	Adg32050 Mutant B_
42	48.5	35.7	594	8	ADG32068	Adg32068 Mutant B_
43	48.5	35.7	616	6	ABR41881	Abr41881 Predicted
44	48.5	35.7	616	6	AAE36781	Aae36781 Human cas
45	48.5	35.7	645	2	AAW29653	Aaw29653 Human sec

ALIGNMENTS

RESULT 1
AAY06970
ID AAY06970 standard; protein; 165 AA.
XX
AC AAY06970;
XX
DT 05-JUL-1999 (first entry)
XX
DE E. canis P30-10protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW detection; dog.
XX
OS Ehrlichia canis.
XX
PN WO9913720-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US019600.
XX
PR 19-SEP-1997; 97US-0059353P.
XX
PA (OHIS) UNIV OHIO STATE.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 1999-254290/21.
DR N-PSDB; AAX34770.
XX
PT Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
PT canis.
XX
PS Disclosure; Fig 30B; 55pp; English.
XX
CC The invention provides isolated outer membrane proteins (OMP) from
CC Ehrlichia chafeensis and E. canis. The E. chaffeensis proteins form part
CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in
CC AAY06943-958. The E. canis proteins form part of the P30 family and
CC consist of proteins shown in AAY06959-970. The proteins and genes are
CC used to detect E. chaffeensis in patients and E. canis in dogs
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 136; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 2
AAU96116
ID AAU96116 standard; protein; 280 AA.
XX
AC AAU96116;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ehrlichia canis p28-2.
XX
KW Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.
XX
OS Ehrlichia canis.
XX
PN WO200222782-A2.
XX
PD 21-MAR-2002.
XX
PF 12-SEP-2001; 2001WO-US028759.
XX
PR 12-SEP-2000; 2000US-00660587.
XX
PA (RERE-) RES DEV FOUND.
XX
PI Walker DH, Yu X, Mcbride JW;
XX
DR WPI; 2002-351882/38.
DR N-PSDB; ABK68876.
XX
PT New recombinant homologous 28 kilodalton immunodominant protein from
PT Ehrlichia canis, useful for treating Ehrlichia canis infections.
XX
PS Claim 16; Fig 14; 106pp; English.
XX
CC The invention relates to a recombinant homologous 28 kDa immunodominant
CC protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably
CC dispersed in a pharmaceutically acceptable carrier, is useful for
CC inhibiting E. canis infection in a subject. (I) is useful in the
CC development of vaccines and serodiagnostics that are particularly
CC effective for disease prevention and serodiagnosis. AAU96100-AAU96118
CC represent the 28-kDa antigen amino acid sequences of the invention
XX
SQ Sequence 280 AA;

```
  Query Match             100.0%;  Score 136;  DB 5;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-12;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 3
ABG77958
ID ABG77958 standard; protein; 280 AA.
XX
AC ABG77958;
XX
DT 15-NOV-2002 (first entry)
XX
DE Ehrlichia canis outer membrane protein (P30F) #9.
XX
KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.
XX
OS Ehrlichia canis.
XX
PN US2002120115-A1.
XX
PD 29-AUG-2002.
XX
PF 28-JAN-2002; 2002US-00059964.
XX
PR 19-MAY-1999; 99US-00314701.
XX
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2002-618954/66.
DR N-PSDB; ABS63299.
XX
PT Isolated polynucleotide encoding an outer membrane protein of E.canis or
PT E.chaffeensis used in the diagnosis of infection.
XX
PS Claim 10; Fig 30B; 49pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding an outer
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
CC in the diagnosis of infection. An infection such as human ehrlichiosis or
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC patient, providing a polypeptide or mixture of polypeptides, contacting
CC the sample with the polypeptide and assaying for the formation of a
CC complex between antibodies in the serum sample and the polypeptide, where
CC formation of a complex is indicative of infection with E. chaffeensis.
CC This sequence represents an Ehrlichia outer membrane protein of the
CC invention
XX
SQ Sequence 280 AA;

```
  Query Match             100.0%;  Score 136;  DB 5;  Length 280;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-12;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              ||||||||||||||||||||||||||
Db         61 PINGTNSLTKKVFGLKKDGDITKKDD 86
```

RESULT 4
ADA09781
ID ADA09781 standard; protein; 280 AA.
XX
AC ADA09781;
XX
DT 06-NOV-2003 (first entry)
XX
DE E. canis outer membrane protein P30-10.
XX
KW outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;
KW Rocky Mountian spotted fever; canine ehrlichiosis; antigen.
XX
OS Ehrlichia canis.
XX
PN US6544517-B1.
XX
PD 08-APR-2003.
XX
PF 19-MAY-1999; 99US-00314701.
XX
PR 18-SEP-1998; 98US-0100843P.
XX
PA (OHIS) UNIV OHIO STATE RES FOUND.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2003-553952/52.
DR N-PSDB; ADA09780.
XX

PT New isolated polynucleotide encoding outer membrane protein P30 of
PT Ehrlichia canis or its variant or fragment, useful for producing
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for
PT diagnosing and treating ehrlichiosis.
XX
PS Disclosure; Fig 30; 105pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a variant of
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer
CC membrane protein of E. canis , or an antigenic fragment of the E. canis
CC P30 protein, or comprising a sequence which is the complement of
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.
CC canis and E. chaffeensis outer membrane proteins and their encoding
CC nucleic acids. The polynucleotides are useful for producing E. canis or
CC E. chaffeensis outer membrane protein, for designing hybridisation probes
CC for isolating and identifying cDNA and genomic clones encoding the OMP)
CC or its allelic forms, for designing primers for PCR. The polypeptides
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present
CC sequence represents an E. canis outer membrane protein.
XX
SQ Sequence 280 AA;

Query Match 100.0%; Score 136; DB 6; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.8e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 5
ADW04274
ID ADW04274 standard; protein; 280 AA.
XX
AC ADW04274;
XX
DT 24-MAR-2005 (first entry)
XX
DE Ehrlichia canis outer membrane protein (P30F), P30-10.
XX
KW DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;
KW infection; vaccine.
XX
OS Ehrlichia canis.
XX
FH Key Location/Qualifiers
FT Peptide 1. .25
FT /label= Signal_peptide
FT Protein 26. .280
FT /note= "Erlichia canis mature OMP protein"
XX
PN US2004265334-A1.
XX
PD 30-DEC-2004.
XX
PF 29-JUL-2004; 2004US-00901774.
XX
PR 18-SEP-1998; 98US-0100843P.
PR 19-MAY-1999; 99US-00314701.
PR 28-JAN-2002; 2002US-00059964.
XX
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2005-064871/07.
DR N-PSDB; ADW04273.
XX
PT New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia
PT canis or Ehrlichia chaffeensis, useful in preparing a composition for
PT diagnosing or preventing E. canis or E. chaffeensis infection.
XX
PS Claim 10; SEQ ID NO 48; 122pp; English.
XX
CC The invention relates to nucleic acid sequences encoding outer membrane
CC proteins (OMP) of Erlichia chaffeensis (designated as OMP proteins) and
CC Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is
CC useful in preparing a composition for diagnosing, treating or preventing
CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present
CC sequence is the Erlichia canis P30F protein.
XX
SQ Sequence 280 AA;

Query Match 100.0%; Score 136; DB 9; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.8e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITKKDD 26
||||||||||||||||||||||||||
Db 61 PINGTNSLTKKVFGLKKDGDITKKDD 86

RESULT 6
AAY06944
ID AAY06944 standard; protein; 283 AA.
XX
AC AAY06944;
XX
DT 27-AUG-2003 (revised)
DT 05-JUL-1999 (first entry)
XX
DE E. chafeensis OMP-1B protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW detection; dog.
XX
OS Ehrlichia chaffeensis.
XX
PN WO9913720-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US019600.
XX
PR 19-SEP-1997; 97US-0059353P.
XX
PA (OHIS) UNIV OHIO STATE.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 1999-254290/21.
DR N-PSDB; AAX34744.
XX
PT Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
PT canis.
XX
PS Claim 12; Fig 4B; 55pp; English.
XX
CC The invention provides isolated outer membrane proteins (OMP) from
CC Ehrlichia chafeensis and E. canis. The E. chaffeensis proteins form part
CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in
CC AAY06943-958. The E. canis proteins form part of the P30 family and
CC consist of proteins shown in AAY06959-970. The proteins and genes are
CC used to detect E. chaffeensis in patients and E. canis in dogs. (Updated
CC on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 283 AA;

Query Match 71.3%; Score 97; DB 2; Length 283;
Best Local Similarity 78.3%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PINGTNSLTKKVFGLKKDGDITK 23

```
              ||||  |:|||||||||||||  :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 7
AAU96106
ID   AAU96106 standard; protein; 283 AA.
XX
AC   AAU96106;
XX
DT   07-AUG-2003  (revised)
DT   02-JUL-2002  (first entry)
XX
DE   Ehrlichia chafeensis OMP-1B.
XX
KW   Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO200222782-A2.
XX
PD   21-MAR-2002.
XX
PF   12-SEP-2001; 2001WO-US028759.
XX
PR   12-SEP-2000; 2000US-00660587.
XX
PA   (RERE-) RES DEV FOUND.
XX
PI   Walker DH,  Yu X,  Mcbride JW;
XX
DR   WPI; 2002-351882/38.
XX
PT   New recombinant homologous 28 kilodalton immunodominant protein from
PT   Ehrlichia canis, useful for treating Ehrlichia canis infections.
XX
PS   Example 3; Fig 3; 106pp; English.
XX
CC   The invention relates to a recombinant homologous 28 kDa immunodominant
CC   protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably
CC   dispersed in a pharmaceutically acceptable carrier, is useful for
CC   inhibiting E. canis infection in a subject. (I) is useful in the
CC   development of vaccines and serodiagnostics that are particularly
CC   effective for disease prevention and serodiagnosis. AAU96100-AAU96118
CC   represent the 28-kDa antigen amino acid sequences of the invention.
CC   (Updated on 07-AUG-2003 to correct OS field.)
XX
SQ   Sequence 283 AA;

  Query Match             71.3%;  Score 97;  DB 5;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 4.2e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||  :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 8
AAU73413
ID   AAU73413 standard; protein; 283 AA.
XX
AC   AAU73413;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Ehrlichia chaffeensis outer membrane protein P28-14.
XX
KW   Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.
XX
OS   Ehrlichia chaffeensis.
XX
PN   WO200183699-A2.
XX
PD   08-NOV-2001.
XX
PF   01-MAY-2001; 2001WO-US013997.
XX
PR   01-MAY-2000; 2000US-0201035P.
XX
PA   (RERE-) RES DEV FOUND.
XX
PI   Walker DH,  Yu X;
XX
DR   WPI; 2002-066527/09.
XX
PT   Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28
PT   useful as a vaccine against Ehrlichia chaffeensis.
XX
PS   Disclosure; Fig 2; 97pp; English.
XX
CC   The invention relates to isolated and purified 28-kDa outer membrane
CC   proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are
CC   encoded by a 28kDa outer membrane protein multigene family. P28 proteins
CC   are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful
CC   for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia
CC   chaffeensis P28 outer membrane proteins of the invention
XX
SQ   Sequence 283 AA;

  Query Match             71.3%;  Score 97;  DB 5;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 4.2e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||  :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 9
ABG77936
ID   ABG77936 standard; protein; 283 AA.
XX
AC   ABG77936;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Ehrlichia chaffeensis outer membrane protein (OMP) #2.
XX
KW   Outer membrane protein; OMP; P30F; ehrlichiosis; infection.
XX
OS   Ehrlichia chaffeensis.
XX
PN   US2002120115-A1.
XX
PD   29-AUG-2002.
XX
PF   28-JAN-2002; 2002US-00059964.
XX
PR   19-MAY-1999;   99US-00314701.
XX
PA   (RIKI/) RIKIHISA Y.
PA   (OHAS/) OHASHI N.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 2002-618954/66.
DR   N-PSDB; ABS63277.
XX
PT   Isolated polynucleotide encoding an outer membrane protein of E.canis or
PT   E.chaffeensis used in the diagnosis of infection.
XX
PS   Disclosure; Fig 4B; 49pp; English.
XX
CC   The invention relates to an isolated polynucleotide encoding an outer
CC   membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
```

```
CC   in the diagnosis of infection. An infection such as human ehrlichiosis or
CC   canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC   patient, providing a polypeptide or mixture of polypeptides, contacting
CC   the sample with the polypeptide and assaying for the formation of a
CC   complex between antibodies in the serum sample and the polypeptide, where
CC   formation of a complex is indicative of infection with E. chaffeensis.
CC   This sequence represents an Ehrlichia outer membrane protein of the
CC   invention
XX
SQ   Sequence 283 AA;

  Query Match             71.3%;  Score 97;  DB 5;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 4.2e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 10
ADA09737
ID   ADA09737 standard; protein; 283 AA.
XX
AC   ADA09737;
XX
DT   06-NOV-2003  (first entry)
XX
DE   E. chaffeensis outer membrane protein OMP-1B.
XX
KW   outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;
KW   Rocky Mountian spotted fever; canine ehrlichiosis; antigen.
XX
OS   Ehrlichia chaffeensis.
XX
PN   US6544517-B1.
XX
PD   08-APR-2003.
XX
PF   19-MAY-1999;   99US-00314701.
XX
PR   18-SEP-1998;   98US-0100843P.
XX
PA   (OHIS ) UNIV OHIO STATE RES FOUND.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 2003-553952/52.
DR   N-PSDB; ADA09736.
XX
PT   New isolated polynucleotide encoding outer membrane protein P30 of
PT   Ehrlichia canis or its variant or fragment, useful for producing
PT   Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for
PT   diagnosing and treating ehrlichiosis.
XX
PS   Disclosure; Fig 4; 105pp; English.
XX
CC   The invention relates to an isolated polynucleotide encoding a variant of
CC   the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer
CC   membrane protein of E. canis , or an antigenic fragment of the E. canis
CC   P30 protein, or comprising a sequence which is the complement of
CC   nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.
CC   canis and E. chaffeensis outer membrane proteins and their encoding
CC   nucleic acids. The polynucleotides are useful for producing E. canis or
CC   E. chaffeensis outer membrane protein, for designing hybridisation probes
CC   for isolating and identifying cDNA and genomic clones encoding the OMP)
CC   or its allelic forms, for designing primers for PCR. The polypeptides
CC   encoded by the polynucleotide is useful for diagnosing human ehrlichiosis
CC   (Rocky mountain spotted fever) or canine ehrlichiosis. The present
CC   sequence represents an E. chaffeensis outer membrane protein.
XX
SQ   Sequence 283 AA;

  Query Match             71.3%;  Score 97;  DB 6;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 4.2e-06;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 11
ADW04230
ID   ADW04230 standard; protein; 283 AA.
XX
AC   ADW04230;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Erlichia chaffeensis outer membrane protein (OMP), OMP-1B.
XX
KW   DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;
KW   infection; vaccine.
XX
OS   Ehrlichia chaffeensis.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .25
FT                   /label= Signal_peptide
FT   Protein         26. .283
FT                   /note= "Erlichia chaffeensis mature OMP protein"
FT   Region          26. .41
FT                   /note= "Semivariable region (SV)"
FT   Region          82. .94
FT                   /note= "Hypervariable region (HV1)"
FT   Region          145. .163
FT                   /note= "Hypervariable region (HV2)"
FT   Region          248. .272
FT                   /note= "Hypervariable region (HV3)"
XX
PN   US2004265334-A1.
XX
PD   30-DEC-2004.
XX
PF   29-JUL-2004; 2004US-00901774.
XX
PR   18-SEP-1998;   98US-0100843P.
PR   19-MAY-1999;   99US-00314701.
PR   28-JAN-2002; 2002US-00059964.
XX
PA   (RIKI/) RIKIHISA Y.
PA   (OHAS/) OHASHI N.
XX
PI   Rikihisa Y,  Ohashi N;
XX
DR   WPI; 2005-064871/07.
DR   N-PSDB; ADW04229.
XX
PT   New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia
PT   canis or Ehrlichia chaffeensis, useful in preparing a composition for
PT   diagnosing or preventing E. canis or E. chaffeensis infection.
XX
PS   Disclosure; SEQ ID NO 4; 122pp; English.
XX
CC   The invention relates to nucleic acid sequences encoding outer membrane
CC   proteins (OMP) of Erlichia chaffeensis (designated as OMP proteins) and
CC   Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is
CC   useful in preparing a composition for diagnosing, treating or preventing
CC   an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present
CC   sequence is the Erlichia chaffeensis OMP protein.
XX
SQ   Sequence 283 AA;

  Query Match             71.3%;  Score 97;  DB 9;  Length 283;
  Best Local Similarity   78.3%;  Pred. No. 4.2e-06;
```

```
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PINGTNSLTKKVFGLKKDGDITK 23
              ||||  |:|||||||||||||| :
Db         64 PINGNTSITKKVFGLKKDGDIAQ 86


RESULT 12
ADN21449
ID   ADN21449 standard; protein; 339 AA.
XX
AC   ADN21449;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Bacterial polypeptide #4102.
XX
KW   Recombinant DNA construct; transformed plant; improved plant property;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW   pathogen tolerance; pest tolerance; plant disease resistance;
KW   cell cycle pathway modification; plant growth regulator;
KW   homologous recombination; seed oil yield; protein yield; carbohydrate;
KW   nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW   bacterial polypeptide.
XX
OS   Bacteria.
XX
PN   US2003233675-A1.
XX
PD   18-DEC-2003.
XX
PF   20-FEB-2003; 2003US-00369493.
XX
PR   21-FEB-2002; 2002US-0360039P.
XX
PA   (CAOY/) CAO Y.
PA   (HINK/) HINKLE G J.
PA   (SLAT/) SLATER S C.
PA   (CHEN/) CHEN X.
PA   (GOLD/) GOLDMAN B S.
XX
PI   Cao Y,  Hinkle GJ,  Slater SC,  Chen X,  Goldman BS;
XX
DR   WPI; 2004-061375/06.
XX
PT   New recombinant DNA construct comprising a promoter positioned to provide
PT   for expression of a polynucleotide encoding a polypeptide from a
PT   microbial source, useful for producing plants with improved properties.
XX
PS   Claim 1; SEQ ID NO 4102; 122pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in a plant cell, where the promoter is positioned to
CC   provide for expression of a polynucleotide encoding a polypeptide from a
CC   microbial source. The invention also relates to a transformed plant
CC   comprising the recombinant DNA construct and a method of producing a
CC   transformed plant having an improved property. The plant is a crop plant
CC   such as maize or soybean. The method of producing a transformed plant
CC   having an improved property comprises transforming a plant with the
CC   recombinant DNA construct and growing the transformed plant, where the
CC   polynucleotide or polypeptide is useful for improving plant properties.
CC   The recombinant DNA construct is useful for producing plants with
CC   improved plant properties, e.g. improved cold, heat or drought tolerance,
CC   tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC   increased resistance to plant disease, better growth rate by modification
CC   of the cell cycle pathway with plant growth regulators, increased rate of
CC   homologous recombination, modified seed oil or protein yield and/or
CC   content, improved yield by modification of carbohydrate, nitrogen or
CC   phosphorus use and/or uptake, by modification of photosynthesis or by
CC   providing improved plant growth and development under at least one stress
CC   condition, improved lignin production or improved galactomannan
CC   production. This sequence represents a bacterial polypeptide used in the
CC   scope of the invention. Note: The sequence data for this patent did not
CC   form part of the printed specification but was obtained in electronic
CC   format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 339 AA;

  Query Match             38.6%;  Score 52.5;  DB 8;  Length 339;
  Best Local Similarity   44.4%;  Pred. No. 31;
  Matches   12;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 PINGTNSLTKKVF-GLKKDGDITKKDD 26
              |:| : |: |||  ||:|:|: |:  |
Db        198 PLNCSESVWKKVLQGLEKEGEGTRDKD 224


RESULT 13
ABU29155
ID   ABU29155 standard; protein; 377 AA.
XX
AC   ABU29155;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #14682.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA33025.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 57079; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
```

```
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 377 AA;

  Query Match             38.2%;  Score 52;  DB 6;  Length 377;
  Best Local Similarity   39.1%;  Pred. No. 43;
  Matches    9;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 NGTNSLTKKVFGLKKDGDITKKD 25
              ||| |:|  :|  |::|::   |
Db        279 NGTKSITASLFKYKENGEVLLND 301


RESULT 14
ADH88024
ID   ADH88024 standard; protein; 378 AA.
XX
AC   ADH88024;
XX
DT   22-APR-2004  (first entry)
XX
DE   Enterococcus faecalis polypeptide #2504.
XX
KW   Enterococcus faecalis infection; transcription regulatory element;
KW   antibacterial.
XX
OS   Enterococcus faecalis.
XX
PN   US6617156-B1.
XX
PD   09-SEP-2003.
XX
PF   13-AUG-1998;   98US-00134000.
XX
PR   15-AUG-1997;   97US-0055778P.
XX
PA   (DOUC/) DOUCETTE-STAMM L A.
PA   (BUSH/) BUSH D.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2003-895394/82.
DR   N-PSDB; ADH84619.
XX
PT   New nucleic acid comprising a sequence encoding an Enterococcus fecalis
PT   polypeptide, useful for preparing a composition for diagnosing or
PT   treating E. fecalis infection.
XX
PS   Disclosure; SEQ ID NO 5909; 193pp; English.
XX
CC   The invention relates to Enterococcus faecalis polynucleotides and
CC   polypeptides. The invention also relates to a recombinant expression
CC   vector comprising a polynucleotide operably linked to a transcription
CC   regulatory element, a cell comprising a recombinant vector, a method for
CC   producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC   a sequence not given in the specification, a recombinant vector
CC   comprising the nucleic acid and a cell comprising the recombinant vector.
CC   The polynucleotides can be used to detect the presence of E. faecalis in
CC   a sample. The sequences are useful for preparing a composition for
CC   diagnosing or treating Enterococcus faecalis infection. This sequence
CC   represents an E. faecalis polypeptide of the invention.
XX
SQ   Sequence 378 AA;

  Query Match             38.2%;  Score 52;  DB 7;  Length 378;
  Best Local Similarity   39.1%;  Pred. No. 43;
  Matches    9;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 NGTNSLTKKVFGLKKDGDITKKD 25
              ||| |:|  :|  |::|::   |
Db        280 NGTKSITASLFKYKENGEVLLND 302


RESULT 15
ABU20708
ID   ABU20708 standard; protein; 481 AA.
XX
AC   ABU20708;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #6235.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Bacteroides fragilis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA24578.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 48632; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
```

CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 481 AA;

Query Match 38.2%; Score 52; DB 6; Length 481;
Best Local Similarity 42.3%; Pred. No. 57;
Matches 11; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

```
Qy          1 PINGTNSLTKKVFGLKKDGDITKKDD 26
              |  | | | :|:     | |||  ||
Db        147 PFGGMNDLKEKIIRTPLDPDITFSDD 172
```

Search completed: January 24, 2006, 11:42:47
Job time : 19.6732 secs